

INSIDE THE BLACK BOX: CHALLENGES IN IMPLEMENTATION EVALUATION OF COMMUNITY MENTAL HEALTH CASE MANAGEMENT PROGRAMS

Margaret Gehrs
St. Michael's Hospital and University of Toronto
Toronto, Ontario

Heather Smith Fowler
University of Ottawa
Ottawa, Ontario

Sean B. Rourke
Donald Wasylenki
St. Michael's Hospital and University of Toronto
Toronto, Ontario

Marnie Smith
Canadian Mental Health Association – Ottawa Branch
Ottawa, Ontario

J. Bradley Cousins
University of Ottawa
Ottawa, Ontario

Abstract: Fidelity measurement is an evolving field in mental health case management program evaluation. This article presents an exploratory study in which two separate fidelity measures, the Dartmouth Assertive Community Treatment Scale (DACTS) and the Key Component Profiles (KCP), were used to assess structure and process elements of three mental health case management programs. The programs were studied because they all provided services to seriously mentally ill inner city populations and shared a common context for practice. However, one program followed the Assertive Community Treatment (ACT) model, the other two were Intensive Case Management (ICM) programs, and one of the ICM programs formed a significant partnership with a home care agency for service delivery. The extent to which the DACTS and KCP were able to measure the structure and process similarities and differences of the programs is examined. The results provide information for evalua-

Corresponding author: Margaret Gehrs, Clinical Leader Manager – Community Mental Health Service, St. Michael's Hospital, 17 Cardinal Carter Wing, 30 Bond St., Toronto, ON M5B 1W8; <gehrsm@smh.toronto.on.ca>

tors on the possible strengths and limitations of each fidelity tool in differentiating various elements of the case management models and reinforce the importance of assessing program fidelity from a multi-dimensional perspective.

Résumé: La mesure de la fidélité est un aspect en évolution dans l'évaluation des programmes de gestion de cas en santé mentale. Cet article présente une étude exploratoire dans laquelle deux mesures distinctes de fidélité, soit la *Dartmouth Assertive Community Treatment Scale* (DACTS) et les *Key Component Profiles* (KCP), ont été utilisées pour évaluer la *structure* et les *processus* de trois programmes de gestion de cas de santé mentale. Ces programmes ont été étudiés parce qu'ils fournissaient tous des services à des personnes souffrant d'un trouble grave de santé mentale et résidant dans les quartiers centraux de la ville, et parce qu'ils présentent un contexte de pratique similaire. Toutefois, l'un des programmes suivait le modèle *Assertive Community Treatment* (ACT) tandis que les deux autres étaient des programmes de gestion intensive de cas (GIC); l'un de ces programmes de GIC a conclu un important partenariat avec une agence de soins à domicile pour la prestation des services. L'étude examine dans quelle mesure la DACTS et les KCP ont été capables de mesurer les similarités et les différences dans la structure et les processus. Les résultats ont fourni aux évaluateurs des renseignements sur les forces et les limitations possibles de chaque instrument de fidélité en ce qui a trait à la différenciation des divers éléments des modèles de gestion de cas et soulignent l'importance d'évaluer la fidélité des programmes d'un point de vue multidimensionnel.

Despite recent trends toward results-based program evaluation (e.g., Gauthier et al., 2004; Mayne, 2004) and the use of RCT designs to demonstrate program effectiveness, valid measurement of program fidelity remains a challenge for evaluators (Mowbray, Holter, Teague, & Bybee, 2003), especially in different comparison conditions. Program fidelity may be defined as "the extent to which delivery of an intervention adheres to the protocol or program model originally developed" (Mowbray et al., 2003, p. 315). Measuring program fidelity well is essential to understanding whether program outcome failure is attributable to inadequacies of the program model or failure to implement the model as intended (Chen, 1990). Approaches to the measurement of program fidelity have been described by Scriven (1997) ranging from relatively high fidelity "glass-box" models, which seek to explicate detailed program theories (e.g., Chen, 1990), to comparatively low fidelity "grey-box" orientations, which focus on the identification and ultimate meas-

urement of less well-defined program components (e.g., see Scheirer, 1994).

In the field of case management, the ability to differentiate and measure case management program sub-types remains an ongoing challenge (Rapp, 1998; Ziguras & Stuart, 2000). Although case management is generally seen as a service that coordinates care and provides practical assistance to seriously mentally ill clients in the community, approaches to case management have evolved differently (Schaedle, McGrew, Bond & Epstein, 2002). Of particular interest to researchers, policy makers, and field clinicians are methods to compare, contrast, and measure the unique features and effectiveness of two types of case management that are expected to play an integral role in mental health reform (Ministry of Health and Long Term Care, 1999). They are Assertive Community Treatment (ACT), a model with clear standards which has proliferated in Ontario since receiving priority funding from the provincial government in 1998 (Sapsford, 1998), and Intensive Case Management (ICM), a model with less clearly defined standards (Schaedle & Epstein, 2000) and a longer and more diverse evolution in the province.

The development of fidelity measurement tools for the case management field is a relatively new area of study. According to Mowbray et al. (2003), the development of fidelity tools needs to consider both the structure and process of the programs under study. *Structure* criteria focus on measurement of concrete, objective indicators such as staffing levels and characteristics, case load size, frequency and intensity of contacts, hours of operation, etc. (Orwin, 2000). *Process* criteria can be more subjective and difficult to rate. They focus on measurement of program style, treatment philosophy, staff-client interactions and therapeutic relationships, individualization of treatment, etc. Because measurement of process is more often based on observations, interviews, and multiple data sources, it requires more time and effort, and can therefore be more costly and less reliable (Mowbray et al., 2003). However, we contend that both approaches are essential to understanding a given program and its implementation.

Both the structure and process of case management programs can be shaped by the contexts in which they operate. For example, mandated standards from government funding bodies regarding hours of operation and staffing levels (such as those imposed by the Ontario Ministry of Health on ACT teams) would certainly influence

the program's performance on structure-focused fidelity measures, since these are the basis for the standards. The community context in which a program operates (the needs of the client population for vocational programs or substance abuse treatment, for example) is an important consideration, as are the values orientation of the agency delivering service (e.g., the degree of consumer advocacy for community care that emphasizes recovery and consumers' strengths) (Anthony, 1993; Jacobson & Greenley, 2001; Rapp, 1993) and the staff educational backgrounds and approaches to forming therapeutic relationships with vulnerable populations, and so on. All these can influence how a particular case management program functions ideologically and how it would therefore score on a more process-oriented fidelity measure.

The purpose of this article is to describe an initiative by researchers involved in the Community Mental Health Evaluation Initiative (CMHEI) to contribute new knowledge to the measurement of program fidelity in the case management field by using both structure- and process-driven fidelity measures to compare and contrast program types. The CMHEI is a five-year, multi-site study funded by the Ontario Ministry of Health. It is a unique provincial initiative in that it contributes to an increased understanding of community mental health program effectiveness. The seven studies included in the CMHEI examined three different types of community mental health services and supports: case management, self-help initiatives (family and consumer), and crisis intervention, as well as a methodological study. Programs were evaluated to determine how they differ on whom they serve, how they affect users over time, and their cost effectiveness (Dewa et al., 2002). Two of the study sites used separate randomized clinical trial (RCT) designs to evaluate outcomes of case management programs for seriously mentally ill clients with homelessness and concurrent substance use issues in inner city areas of Toronto and Ottawa.

The case management programs selected for study practice in a similar community context. In particular, all the programs have adapted their practice to work with inner city clients with serious mental illness, homelessness, and substance abuse issues. While all are forms of case management, a major difference in the programs is that one has been funded to follow ACT standards and the other two are ICM programs. The two ICM programs are also different in that one of the programs formed a significant partnership with a home care program, resulting in shared delivery of clinical services

to their clients. Despite these obvious *qualitative* similarities and differences, measurement of *quantitative* similarities and differences for the purpose of program evaluation in the CMHEI RCTs could only be determined using tested fidelity measurement tools that examine both structure and process elements. This article describes an exploratory approach in applying two fidelity measurement tools, the Dartmouth Assertive Community Treatment Scale (DACTS) (Teague, Bond, & Drake, 1998) and the Key Component Profiles (KCP) (Cousins, Aubry, Smith Fowler, & Smith, 2004) to measure structure and process elements of the three case management programs. In particular, we hypothesize that the ACT program will demonstrate higher scores on the more structure-oriented DACTS, and that the more process-oriented ICM programs will score higher on the measure that assesses this in more detail — the KCP. We expect that the results of this study will provide insights into the correlation between program models and outcome results in the CMHEI RCTs and in the case management field in general.

METHOD

Description of the Programs

The following section outlines the Toronto and Ottawa RCT studies and the programs involved in each. Program descriptions provide a brief synopsis of select structure elements of each program and of some of the ideological elements influencing the process of each program. Program characteristics are summarized in Table 1.

The Toronto CMHEI Study: ACT and ICM Programs

The purpose of this RCT study was to investigate the effectiveness of an ACT program and an ICM program, partnered with a home care service, in a Canadian inner city environment. Another study objective was to demonstrate that such programs could be established and would be acceptable to a very high-need client population in a disadvantaged urban environment. The target population for both programs consisted of individuals with serious and persistent mental illness, many of whom are homeless with concurrent substance use disorders, very few resources, and conditions often complicated by chronic medical illnesses. A total of 40 subjects were randomly allocated to each of the study's comparative treatment groups.

The Toronto ACT Program. The Toronto ACT program, part of St. Michael's Hospital Mental Health Service, was implemented to closely model the Ontario Ministry of Health and Long Term Care standards for ACT teams (Gallow, 1997; Sapsford, 1998). Throughout the duration of the study the program operated seven days a week, 24 hours a day, and served a caseload of approximately 80 clients who required assertive outreach in the community. The ACT team consisted of 10 full-time case managers of various disciplines,

Table 1
Characteristics of Toronto and Ottawa Programs

	Toronto ACT	Toronto ICM	Ottawa ICM
Year program established	1996	1998	1986
Sponsoring agency	Teaching hospital	Teaching hospital	Community mental health agency
Case manager full-time equivalents	10	5 to 6	15
Disciplinary designations of case managers	Nursing Occupational Therapy Social Work Addictions Counselling Vocational Rehabilitation Recreational Therapy	Nursing Occupational Therapy Social Work Recreational Therapy	Generic Case Managers
Other team members	Peer Support Worker Psychiatrists Team Leader	Psychiatrist Team Leader	Substance Abuse Specialist
Service delivery partnerships		Significant partnership with homecare services. Nursing and personal support brokered to community agencies.	
Hours of operation	24/7	Case managers: 9–5, Monday to Friday Home care services: available for extended hours seven days per week.	9–5 Monday to Friday
Caseloads	8	15	15
Philosophy	Psychosocial rehabilitation Recovery Risk Management	Recovery	Recovery Strengths-based Psychosocial rehabilitation

including nursing, social work, occupational therapy, vocational rehabilitation, addictions counselling, and recreation therapy. Each client was assigned a primary case manager; however, implementation of service plans was shared by all members of the team. The team also included a part-time peer support worker, who provided a consumer perspective within the team. Two part-time psychiatrists provided both outpatient and inpatient care as needed, thus achieving continuity of care.

Team members described the program as following psychosocial rehabilitation (Cnaan, Blankertz, & Saunders, 1992; Farkas, Anthony, & Cohen, 1989) and recovery principles (Anthony, 1993; Jacobson & Greenley, 2001). Clients were encouraged to have input into their service plans, and there was a strong emphasis on the Individual Placement and Support (IPS) vocational rehabilitation model (Drake, Becker, Clark, & Mueser, 1999), which works with clients' strengths. However, the team admitted that client choice was sometimes compromised when a risk management approach was instituted for clients who were very ill, vulnerable, or incapable of giving consent. In such situations, decisions were made for the client, as warranted, to maintain safety in the community.

The Toronto ICM Program. The Toronto ICM program is also part of St. Michael's Hospital Mental Health Service. Throughout the duration of the study it operated Monday to Friday, 9 a.m. to 5 p.m. each day. The ICM team consisted of a team of five to six case managers with varying professional backgrounds, and one part-time psychiatrist. Each case manager carried an individual caseload of up to 15 clients. The psychiatrist provided direct treatment to most clients; however, flexibility existed in that clients could receive services from other psychiatrists instead. Unlike the ACT team, the psychiatrist did not provide inpatient treatment to the ICM clients.

This ICM program was unique from other ICM programs in the community in that it formed a significant partnership with a home care service, the Toronto Community Care Access Center, for nursing and personal support services. The nursing services, brokered from St. Elizabeth Health Care, and the personal support services, brokered from the Visiting Homemakers' Association, were coordinated by a Toronto Community Care Access Center home care coordinator, who communicated with the ICM team regarding services required. By partnering with these agencies, as well as other crisis supports in the community, support to clients was extended to seven days per week

and after hours. The partnership also offered more flexibility in terms of frequency of visits than could be offered by the ICM case managers alone. For example, clients could be seen daily if required.

The case managers espoused a strong recovery-oriented perspective (Anthony, 1993; Jacobson & Greenley, 2001) focusing heavily on client input and client goals. Partner agencies were educated about the recovery model in a training session, but more focus was placed on attending to concrete needs of clients (i.e., administering medications, assisting with personal care) than articulating a philosophy of care.

The Ottawa CMHEI Study: ICM

The Ottawa study also employed an RCT design, the purpose of which was to evaluate the effectiveness of an intensive case management (ICM) program run by the Ottawa branch of the Canadian Mental Health Association (CMHA). Services are targeted to people with severe and persistent mental illness who are homeless or at risk of becoming homeless. Many of these people have co-occurring substance abuse problems as well as other complex needs, and most have few resources in terms of formal or informal supports. In this study, the comparison or control condition was standard care in the community, with referrals made by housing outreach workers. A total of 147 people participated in the study: 75 in the treatment group and 72 in the control group. For the purposes of this paper, only the ICM program will be discussed.

The Ottawa ICM program. CMHA Ottawa's ICM program provides clinical case management services, which include assertive outreach, assessment, service planning and coordination, counselling, advocacy, and some crisis intervention. The ICM program is based on a one-to-one relationship between case manager and client, with back-up support available from another case manager during holidays and after hours. During the study, hours of service were from 9 a.m. to 5 p.m., Monday to Friday. The caseload ratios were generally 15:1, although this increased slightly depending on the number of staff on leave, turnover, etc. Case managers came from various disciplines (primarily from social work and psychology), and received extensive initial and ongoing training in the principles and practice of case management, as well as in concurrent disorders, assessment, crisis intervention, and specific treatment modalities.

The Ottawa ICM services are client-directed, recovery-oriented (Anthony, 1993; Jacobson & Greenley, 2001) and largely based on a

strengths-based case management model (Rapp, 1993; Rapp & Wintersteen, 1989), with strong psychosocial rehabilitation underpinnings (Farkas et al., 1989). Emphasis is placed on helping clients fulfill basic needs such as housing and access to services, as well as broader life goals such as educational, vocational, and relationship-oriented goals.

Measures

The Dartmouth Assertive Community Treatment Scale (DACTS)

Structure evaluation of the case management programs was accomplished through the use of the Dartmouth Assertive Community Treatment Scale (DACTS). The DACTS is a 28-item, interviewer-administered instrument developed by Teague et al. (1998), and cited by Salyers et al. (2003) as the most systematic approach to assessing fidelity to the ACT model. Like the Index of Fidelity of Assertive Community Treatment (IFACT) (McGrew, Bond, Dietzen, & Salyers, 1994), it is a rating scale of model characteristics at the program level, and captures information about service delivery, especially those related to structure elements (Calsyn, 2000). Key dimensions include human resources (e.g., staffing complement, caseloads), organizational boundaries (e.g., intake rate, responsibility for crisis and hospital admissions/discharges), and the nature of services (e.g., frequency and intensity of services). Graduated scoring of elements is based on an extensive review of program documents and some primary data collection with staff and, in some cases, clients. There is evidence that it has the potential to discriminate the structure elements of ACT from other types of case management programs (Teague et al., 1998), and that high scores may be correlated with improved client outcomes (McHugo, Drake, Teague, & Xie, 1999).

Although the DACTS was primarily developed to measure the fidelity of ACT programs, researchers have more recently applied the instrument to ICM and brokered case management programs as a means of establishing score norms for these models. Norm-referenced interpretations of fidelity can then be made by comparing scores of individual programs to scores of large, representative groups of same-type programs (Salyers et al., 2003). Although a limitation of this approach is that ACT-driven fidelity criteria may not capture all the necessary components of other case management models, it does allow researchers to evaluate how various models fall along the continuum of program structure elements included in this scale.

Key Component Profile (KCP)

The KCP instrument was developed by researchers from the Ottawa CMHEI study team (Cousins et al., 2004) in order to have a more relevant mechanism for measuring implementation of the Ottawa ICM program. This instrument is of particular interest in that it focuses on measuring characteristics of case management services at the client or case level, as opposed to the program structure level, and hence has the potential to identify in a more specific manner the links between process and outcome. The development of the KCP was grounded in Leithwood and Montgomery's (1987) work on innovation profiles to measure educational innovation, in which multiple dimensions of a given curriculum innovation were differentiated into qualitatively distinct levels of performance, ranging from typical to exemplary. In adapting the approach to ICM, the Ottawa team used a detailed, multi-step process. This included a review of program logic and the activity descriptions of case managers, focus groups with case managers, a pilot study, and item analysis to identify key process components of the ICM program and their five behavioural dimensions. These dimensions are: (a) helping clients meet their basic needs, (b) helping to develop and coordinate a network of formal supports, (c) helping clients with their informal network, (d) facilitating access to services, and (e) facilitating personal goal achievement. Each dimension has three or four levels of implementation, ranging from minimal to full or complete implementation, ranging from "0" (minimal implementation) to "1" (full implementation). The implementation rating was weighted by multiplying it by an importance rating, ranging from 1 to 5, reflecting the priority given to a particular dimension in the services provided by the worker within the given time period. While the measure is new and in need of further testing, initial reliability and validity data are reasonable (see Cousins et al., 2004).

This study represents the first time the KCP has been applied to another ICM program and to an ACT program serving similar populations, for the purpose of comparing process elements at the case manager-client level across programs.

Procedure

Use of DACTS in Toronto and Ottawa

The Toronto and Ottawa sites each used independent raters (members of the respective research teams) to score their programs using

the DACTS score sheet. In both cases, the process included a combination of interviews with program managers and reviews of clients' charts and program documents. During the course of the DACTS scoring, the raters consulted each other periodically to ensure that data collection protocols were as consistent as possible. For example, since the studies at both locations continued for some time, it was decided that a rating date roughly halfway through the recruitment phase of participants at each site would be more representative than a fixed date. Scoring was guided by the Protocol for Assertive Community Treatment Fidelity Scale (Dartmouth Assertive Community Treatment Scale — DACTS) (Teague et al., 1998).

A summary of DACTS scores for the Toronto and Ottawa ACT and ICM programs is provided in Table 2. In order to determine the degree to which the home care partnership affected the structure of the Toronto ICM program, two sets of scores were provided in that case: one for the ICM program alone and one for the ICM program plus home care services. The same rater completed the second DACTS score sheet, using additional data available from the Toronto Community Care Access Centre records.

Use of KCP in Ottawa and Toronto

Once developed, the KCP measure was used extensively by the Ottawa site to assess program implementation at the same intervals as outcomes were assessed, starting at 9 months and continuing to 18 and 24 months. A KCP was completed by a case manager to reflect the previous month of service for each ICM client participating in the study. Ratings were completed for 51 ICM clients.

The Toronto case managers completed their ratings of program implementation in August 2003. In order to assist with this process, the coordinator of the Ottawa project provided the same background information on the measure and instructions for its use as was generally used with Ottawa staff to the Toronto researcher who coordinated data collection at that end. He in turn provided that information to each case manager (of both ACT and ICM programs) on a one-to-one basis. KCP ratings were completed for 10 clients from each program. These clients were selected on the basis that they had participated in the randomized controlled trial and had maintained a relationship with the case manager completing the scale throughout the duration of the study, that is, at least 18 months. The 18-month data from the Ottawa ICM program was compared to the Toronto data collected in August 2003.

Table 2
DACTS 28 Item Scores for Individual Programs

Criteria	Ratings			
	Toronto ACT	Toronto ICM Only*	Toronto ICM + Home Care Services*	Ottawa ICM
Human Resources Criteria				
H1 Small caseload	5	4	5	4
H2 Team approach	5	3	5	1
H3 Program meeting	4	3	3	2
H4 Practicing team leader	1	5	5	1
H5 Continuity of staffing	4	4	4	4
H6 Staff capacity	5	5	5	3
H7 Psychiatrist on staff	5	2	2	1
H8 Nurse on staff	5	5	5	1
H9 Substance abuse specialist on staff	5	1	1	1
H10 Vocational specialist	4	1	1	1
H11 Program size	5	3	4	5
Human Resources Mean	4.36	3.27	3.64	2.18
Organizational Boundaries Mean				
O1 Explicit admission criteria	5	5	5	5
O2 Intake rate	5	5	5	5
O3 Full responsibility for treatment services	5	3	4	2
O4 Responsibility for crisis services	5	1	2	1
O5 Responsibility/ hospital admissions	5	2	2	4
O6 Responsibility for hospital discharge planning	5	2	2	3
O7 Time-unlimited services	5	3	3	5
Organizational Boundaries Mean	5.00	3.00	3.28	3.57
Nature of Services				
S1 Community-based	5	5	5	5
S2 No drop-out policy	5	4	4	5
S3 Assertive engagement	5	4	4	4
S4 Intensity of service	4	2	4	3
S5 Frequency of contact	3	1	5	2
S6 Work with informal supports	3	3	3	3
S7 Individualized substance abuse	4	3	3	3
S8 Dual disorder treatment groups	2	1	1	1
S9 Dual disorders model	5	5	5	3
S10 Role of consumers	3	2	2	5
Nature of Services Mean	3.9	3.0	3.6	3.4
Mean Total	4.42	3.09	3.51	3.05

*The same program is rated twice; in one case the structure of only the ICM program is rated and in the other case the structure of the ICM program plus Home Care Services is rated.

RESULTS

To test the first hypothesis, a norm-referenced interpretation of DACT scores for the three programs was used. This involved comparing the DACT scores for each program to pre-established, published norms for 27 ACT programs in New York State to determine how similar the three programs were to structural norms of ACT programs. The DACTS scores for the Toronto and Ottawa ICM programs were also matched against published norms for 25 ICM programs tested in the United States, to determine how much structure similarity they had with other ICM programs. Although the 28-item DACTS was used in this study, only 26 items could be matched against the published norms. This is due to the fact that the DACTS had evolved from a 26- to a 28-item scale over time and published norms were only available for 26 items. Those items eliminated for analysis centred around program staff size and the role of consumers on the team (Salyers et al., 2003).

Standardized scores were calculated for each item [$Z_{1-26} = (\text{Mean Salyers item}_{1-26} - \text{ACT/ICM item}_{1-26}) / \text{Standard Deviation item}_{1-26}$]. Scores that were within one standard deviation from the mean of the reference group fell within the norm for the group; scores more than one standard deviation from the mean were outside the norm of the group. The number of items that fell inside and outside the norms of the reference groups was determined for each program in order to gain a general picture of the structure similarities and differences in the programs. It should be noted that the standard deviations for the programs ranged significantly, evidence that the comparison programs also varied widely on this measure of program fidelity.

To test the second hypothesis, KCP scores for the Toronto and Ottawa programs were first weighted by the importance rating each case manager had given that element to his or her ongoing work with the individual client, and then averaged, to yield a weighted mean for each of the five dimensions. The scores for each program were then averaged, to yield an overall mean for each dimension by program. A one-way ANOVA was conducted to assess the degree of variance among these weighted means for the three programs, and a post-hoc analysis was conducted to determine the specific relationships (using Tukey's HSD adjustment where equal variances were demonstrated, and Games-Howell for Dimension 2, where variances were not equal).

DACTS Scores: Between Program Comparisons

When the DACT scores for all three Toronto and Ottawa programs were plotted on a norm-referenced graph (see Figures 1 and 2), obvious differences in the structural elements of all three programs emerged. Figure 1 compares the study programs to the norm-referenced data elicited from 27 ACT teams in New York State (Salysers et al., 2003), indicated by the shaded area of the graph. Of the 26 structure items measured, the Toronto ACT program's structure fell above the ACT team norms on 4 items, within the norms on 20 items, and below the norms on 2 items, indicating a high fidelity to structure components of the ACT model as compared to the reference group.

The Toronto and Ottawa ICM programs demonstrated similar patterns when comparing their ratings on the 26 items. Although they did not share the same level of compliance with the ACT norms as was the case with the Toronto ACT program (as was expected given the different structural nature of the programs), they did show a surprising number of structure elements in common with the ACT model, particularly in terms of the Nature of Services dimension and the presence of specialized treatment staff. Both programs had 15 items that fell within the norm for the ACT teams and 11 that were outside the norms. Structure differences, however, did exist between the two ICM programs. For example, the Toronto ICM program had one item that actually surpassed the ACT team norms (S9 - Integration of dual disorder model). There was also variation on which items fell inside and outside the ACT norms for each ICM program and how many standard deviations from the mean each program scored on each item.

Of interest, however, were the scores of the Toronto ICM combined with its home care service partner. By including the impact of the partnership in the analysis of the Toronto ICM program structure, it was noted that the number of items falling within the ACT norms shifted from 15 to 19. More important, three critical program structure elements (H1 - Caseload size, S4 - Intensity of service, and S5 - Frequency of contact) shifted from measuring significantly below the ACT mean to meeting the ACT mean levels for case load size and intensity of service and exceeding the ACT mean for frequency of contact. Similarly, the Toronto ICM partnership model seemed to match the Toronto ACT program in the mean caseload size and intensity of service, but measured significantly higher in mean frequency of contact.

Figure 1
Norm-Referenced Comparison of ACT and ICM Programs in Toronto and Ottawa on the Dartmouth Assertive Community Treatment Scale (Salyers et al., 2003; New York ACT ($n = 27$))

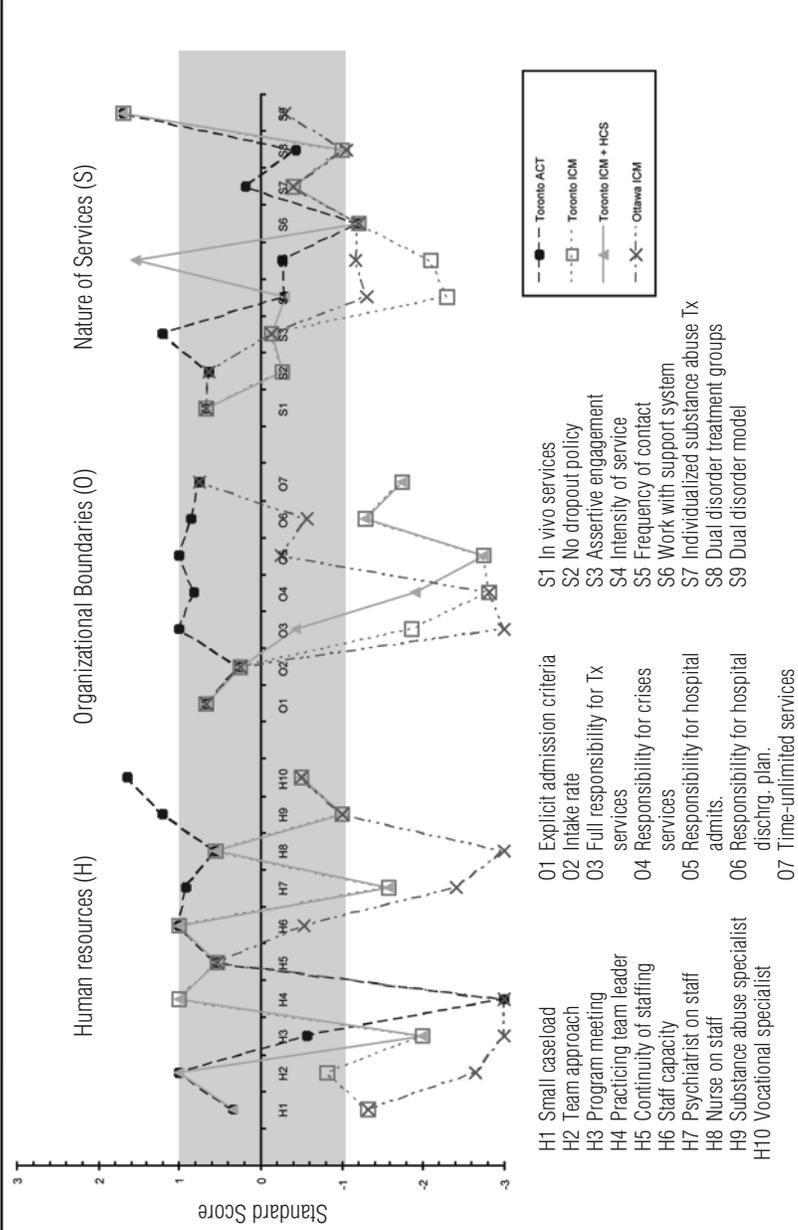
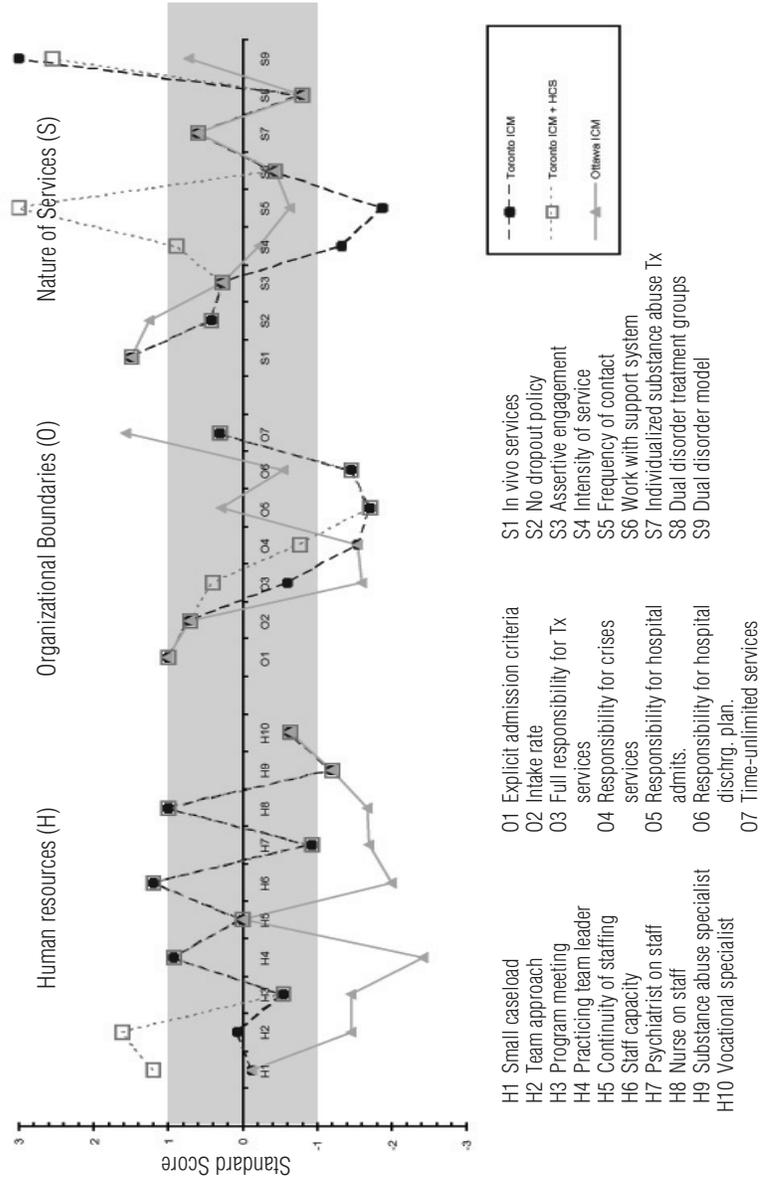


Figure 2
Norm-Referenced Comparison of Toronto and Ottawa ICM Programs on the Dartmouth Assertive Community Treatment Scale (Salyers et al., 2003; ICM (n = 25))



- H1 Small caseload
- H2 Team approach
- H3 Program meeting
- H4 Practicing team leader
- H5 Continuity of staffing
- H6 Staff capacity
- H7 Psychiatrist on staff
- H8 Nurse on staff
- H9 Substance abuse specialist
- H10 Vocational specialist
- O1 Explicit admission criteria
- O2 Intake rate
- O3 Full responsibility for Tx services
- O4 Responsibility for crises services
- O5 Responsibility for hospital admits.
- O6 Responsibility for hospital dischrg. plan.
- O7 Time-unlimited services
- S1 In vivo services
- S2 No dropout policy
- S3 Assertive engagement
- S4 Intensity of service
- S5 Frequency of contact
- S6 Work with support system
- S7 Individualized substance abuse Tx
- S8 Dual disorder treatment groups
- S9 Dual disorder model

Figure 2 compares the degree of similarity in DACTS scores between the Toronto and Ottawa ICM programs and the reference group of 25 American ICM programs (Salyers et al., 2003). Compared to other ICM programs being measured against ACT structure elements, the Toronto ICM program demonstrated 3 of the 26 items above the ICM norms, 17 within the ICM norm, and 6 below the ICM norm. The Ottawa program scored 3 of 26 items above the ICM norm, 14 within the ICM norm, and 9 below the ICM norm.

Of interest again is the impact of the home care service partnership on the Toronto ICM structure scores. By factoring in the home care service partnership, the Toronto ICM mean for the item “S4 - Intensity of service” shifted from below the ICM norm to within the ICM norm, whereas the item “S5 - Frequency of contact” shifted from below the ICM norm to significantly above the ICM norm. The item “H1 - Small caseload” shifted from within the ICM norm to above the ICM norm.

KCP Scores: Between Program Comparisons

In terms of the KCP analysis, the Toronto ICM program scored higher than either the Ottawa ICM program or the Toronto ACT program on every one of the five dimensions (see Table 3). There was no pattern to the scores of the latter two programs relative to each other. The Ottawa ICM program scored higher than the Toronto ACT program on Dimensions 3 and 5 relating to informal supports and goal achievement respectively, and marginally higher on Dimension 1 relating to basic needs. The Toronto ACT program scored higher than the Ottawa ICM program on Dimensions 2 and 4 relating to access to formal supports and rights and responsibilities.

An analysis of variance found statistically significant differences among the groups on Dimension 1 ($F = 3.11, p = .05$) and Dimension 5 ($F = 8.46, p = .001$). Differences on Dimension 2 ($F = 2.86, p = .06$) and Dimension 3 ($F = 2.56, p = .09$) approached significance. Post-hoc analyses indicated that the difference on Dimension 1 was significant between the two ICM programs (Tukey's HSD $p = .04$), whereby the Toronto program was rated as having a higher level of implementation than the Ottawa program on helping clients fulfill basic needs. Other post-hoc comparisons found no differences between programs on Dimension 1. Post-hoc comparisons found differences among all three programs on Dimension 5 (i.e., Bonferroni test indicating significance at or below $p = .002$). These differences

were in the direction of the Toronto ICM being rated the highest on implementation in the area of goal attainment, followed by the Ottawa ICM program, and then the Toronto ACT program.

DISCUSSION

The first hypothesis was supported in that the DACTS was indeed able to demonstrate structural differences between the Toronto ACT

Table 3
Key Component Profile Scores for Ottawa and Toronto

		<i>N</i>	Mean	Minimum	Maximum	Standard Deviation
Basic Needs						
Items 1-8	Ottawa ICM	51	2.84	.70	4.50	0.85
Weighted item mean	Toronto ACT	10	2.81	1.07	4.50	1.10
	Toronto ICM	10	3.59	2.34	5.00	0.84
	Total	71	2.94	0.70	5.00	0.95
Formal Supports						
Items 1-3	Ottawa ICM	51	2.15	0.58	4.00	0.95
Weighted item mean	Toronto ACT	10	2.68	1.17	4.58	1.16
	Toronto ICM	10	2.95	0.50	4.67	1.61
	Total	71	2.34	0.50	4.67	1.11
Informal Supports						
Items 1-3	Ottawa ICM	51	2.03	0.28	5.00	1.27
Weighted item mean	Toronto ACT	10	1.45	0.47	2.83	0.76
	Toronto ICM	10	2.69	0.94	4.33	1.25
	Total	71	2.04	1.25	0.94	4.33
Rights and Responsibilities						
Items 1-2	Ottawa ICM	51	2.52	.42	5.00	1.38
Weighted item mean	Toronto ACT	10	2.67	.46	5.00	1.58
	Toronto ICM	10	3.36	.42	5.00	1.95
	Total	71	2.66	.42	5.00	1.50
Goal Achievement						
Items 1-6	Ottawa ICM	51	2.21	.42	4.83	1.04
Weighted item mean	Toronto ACT	10	1.94	.46	3.67	1.04
	Toronto ICM	10	3.60	1.71	4.79	1.01
	Total	71	2.36	.42	4.83	1.14

and the Toronto and Ottawa ICM programs. More surprising, however, was that it was also able to demonstrate structural differences between the Toronto and Ottawa ICM programs and between these programs and the ICM programs cited in the literature (Salyers et al., 2003). The fact that different ICM program structures can evolve to serve very similar populations in very similar community contexts, as in the Toronto and Ottawa examples, demonstrates once again the diverse evolution of ICM programs in the field (Schaedle et al., 2002). The fact that both the Toronto and Ottawa ICM programs showed differences from ICM structural norms of the American sample also reinforces the fact that no clear structure standards exist for the ICM model. However, further analysis of DACTS scores may be useful in determining where convergence in structural elements is evolving.

An important revelation in the analysis of the DACTS data is the significant contribution that partner agencies can make to the structural nature of case management services delivered to clients in the community. For example, on its own, the Toronto ICM program was not able to deliver the same intensity nor frequency of service that was possible when the home care partner agencies assisted in delivering the needed services. In fact, when working in partnership, the Toronto ICM and home care service were actually able to exceed ACT norms in frequency of service, thereby exceeding the “dosage” or amount of treatment (Calsyn, 2000) provided by all the other programs.

The impact of other services, more often delivered through informal than formal partnerships with case managers in the community, may have a significant impact on the nature of actual services received by a client. In conjunction with case management services, other services such as housing support services, peer and family support, vocational and social-recreational agencies, and so on could significantly enhance the intensity and frequency of support actually received by a client. When examining program models in outcome studies, researchers may need to investigate the larger network of supports available to clients in a case management program, in addition to the program’s own structure, to determine the true complexity of services delivered. The availability of additional services may vary greatly between communities, thereby impacting the structure of the ICM, or even ACT, model in various geographic settings.

The second hypothesis was only partially supported, in that the KCP was able to differentiate some process elements among the three

programs, but not necessarily in the expected directions. In particular, it revealed significant differences among all three programs in terms of how they implemented services related to clients achieving goals in broad life domain areas (Dimension 5). Given that the KCP was developed on the basis of the program logic and activity descriptions of an ICM program, it is not surprising that the Toronto ICM program would score high. What is surprising, however, is that the scores for this program should be so much higher than that of the other ICM program, and that the Toronto ACT program did not consistently score lower than the Ottawa ICM program on every dimension, indicating that some of the key elements of its service may not, in fact, be so different from those of the Ottawa ICM program.

These findings may be due to the limitations of the small sample size for the Toronto programs. It is also possible that more reliability study of the KCP tool is necessary to determine its applicability across different programs. However, it is also possible that the processes inherent in all three programs were indeed more similar than different, despite the ACT-ICM dichotomy. All three programs served very similar populations: homeless seriously mentally ill clients with substance use problems. These programs may attract similar personnel who are drawn to this population and choose to work from a recovery- and strengths-based perspective, as the qualitative descriptions of the programs seemed to indicate. As principles of recovery become more widespread in the case management literature, there may be a greater-than-anticipated convergence of ideology guiding process across case management models. More widespread application of the KCP across case management programs and models, in different community contexts and with different sub-populations, may be necessary to determine its usefulness as a fidelity measurement tool to test the ICM model, or simply to assess fidelity to strengths-based or recovery-based processes in general.

CONCLUSION

Although the ACT model is often seen as a glass-box condition because of the extensive fidelity testing of the model's structure, more investigation is required to determine if the process elements of this model are more similar or different from the ICM, or other case management models. This is especially true as recovery- and strengths-based ideology expands across the mental health service system. To further explicate the grey-box condition of ICM, both

structure and process elements require further investigation. To address this challenge, researchers will need to develop fidelity measures that assess multi-dimensional aspects of the ICM and other case management models. Although the DACTS has proven to be a valid tool in measuring structure elements of ACT and other case management models, it was not designed to measure the more elusive therapeutic processes that occur at the case manager-client level and that may influence client outcomes (Calsyn, 2000). Similarly, the KCP demonstrates strengths in gathering process-oriented data regarding case management programs, but its ability to differentiate among programs is still in question. Regardless of the limitations of these tools, they are helpful in assisting researchers to contribute to the existing literature (Schaedle & Epstein, 2000; Schaedle et al., 2002), by determining where structure and process convergence is occurring across ICM programs and how this model compares to ACT. This is helpful, not only for researchers engaging in RCTs and program evaluation of these models, but also for policy makers who strive to understand the unique role of these models in promoting positive outcomes in mental health reform.

In the further development of fidelity measurement processes, researchers must also find ways to further evaluate the context in which case management programs operate. The context may influence the number of informal and formal support that are available in a community to boost the services offered by case management programs, thereby influencing client outcomes. The context may also shape the ideology and processes of case managers, more than the model label itself.

ACKNOWLEDGEMENTS

The authors acknowledge the important contributions of Michael Higgins, Research Coordinator, St. Michael's Hospital, in collecting data for this project. The authors also thank the staff of the following programs for their willing cooperation and assistance with the research:

- CONTACT Mental Health Outreach Service and the Community Connections Case Management Program, St. Michael's Hospital
- Toronto Community Care Access Centre
- Intensive Case Management and Outreach Programs, Canadian Mental Health Association—Ottawa Branch.

REFERENCES

- Anthony, W. (1993). Recovery from mental illness: The guiding vision of the mental health service for the 1990s. *Psychosocial Rehabilitation Journal*, 16(4), 11–23.
- Calsyn, R.J. (2000). A checklist for critiquing treatment fidelity studies. *Mental Health Services Research*, 2(2), 107–113.
- Chen, H.T. (1990). *Theory-driven evaluations*. Thousand Oaks, CA: Sage.
- Cnaan, R.A., Blankertz, L., & Saunders, M. (1992). Perceptions of consumers, practitioners, and experts regarding psychosocial rehabilitation principles. *Psychosocial Rehabilitation Journal*, 16(1), 95–119.
- Cousins, J.B., Aubry, T.D., Smith Fowler, H., & Smith, M. (2004). Using key component profiles for the evaluation of program implementation in intensive mental health case management. *Evaluation and Program Planning*, 27, 1–23.
- Dewa, C.S., Durbin, J., Wasylenki, D., Ochoka, J., Eastabrook, S., Boydell, K.M., & Goering, P. (2002). Considering a multisite study? How to take the leap and have a soft landing. *Journal of Community Psychology*, 30, 1–15.
- Drake, R.E., Becker, D., Clark, R.E., & Mueser, K.T. (1999). Research on the individual placement and support model of supported employment. *Psychiatric Quarterly*, 70(4), 289–301.
- Farkas, M.D., Anthony, W.A., & Cohen, M.R. (1989). Psychiatric rehabilitation: The approach and its problems. In M.D. Farkas & W.A. Anthony (Eds.), *Psychiatric rehabilitation programs: Putting theory into practice* (pp. 1–27). Baltimore: Johns Hopkins University Press.
- Gallow, M. (1997). *Assertive Community Treatment team guideline*. Toronto: Ministry of Health.
- Gauthier, B., Barrington, G., Bozzo, S. L., Chaytor, C., Cullen, J., Lahey, R., et al. (2004). The lay of the land: Evaluation practice in Canada today. *Canadian Journal of Program Evaluation*, 19(1), 143–178.
- Jacobson, N., & Greenley, D. (2001). What is recovery? A conceptual model and explication. *Psychiatric Services*, 52(4), 482–485.

- Leithwood, K.A., & Montgomery, D.J. (1987). *Improving classroom practice using innovation profiles*. Toronto: Ontario Institute for Studies in Education Press.
- Mayne, J. (2004). Reporting on outcomes: Setting performance expectations and telling performance stories. *Canadian Journal of Program Evaluation*, 19(1), 31–60.
- McGrew, J. H., Bond, G. R., Dietzen, L., & Salyers, M. (1994). Measuring the fidelity of implementation of a mental health program model. *Journal of Counselling and Clinical Psychology*, 62, 670–687.
- McHugo, G.J., Drake, R.E., Teague, G.B., & Xie, H. (1999). The relationship between model fidelity and client outcomes in the New Hampshire Dual Disorders Study. *Psychiatric Services*, 50, 818–824.
- Ministry of Health and Long Term Care. (1999). *Making it happen: Implementation plan for mental health reform* (Cat. No. 2231504). Toronto: Queen's Printer for Ontario.
- Mowbray, C.T., Holter, M.C., Teague, G.B., & Bybee, D. (2003) Fidelity criteria: Development, measurement and validation. *American Journal of Evaluation*, 24, 315–340.
- Orwin, R.G. (2000). Assessing program fidelity in substance abuse health services research. *Addiction*, 95(Suppl.3), S309–S327.
- Rapp, C.A. (1993). Theory, principles and methods of the strengths model of case management. In M. Harris & H.C. Bergman (Eds.), *Case management for mentally ill patients: Theory and practice* (pp. 143–164). Langhorne, PA: Harwood Academic Publishers.
- Rapp, C.A. (1998). The active ingredients of effective case management: A research synthesis. *Community Mental Health Journal*, 34, 363–380.
- Rapp, C.A., & Wintersteen, R. (1989). The strengths model of case management: Results of twelve demonstrations. *Psychological Rehabilitation*, 1, 23–32.
- Salyers, M.P., Bond, G.R., Teague, G.B., Cox, J.F., Smith, M.E., Hicks, M.L., & Koop, J.I. (2003). Is it ACT yet? Real-world examples of evaluating the degree of implementation for assertive community treatment. *Journal of Behavioral Health Services and Research*, 30, 304–320.

- Sapsford, R. (1998). *Standards for Assertive Community Treatment Teams (ACTT)*. Toronto: Ministry of Health.
- Schaedle, R.W., & Epstein, I. (2000). Specifying intensive case management: A multiple perspective approach. *Mental Health Services Research, 2*, 95–105.
- Schaedle, R., McGrew, J.H., Bond, G., & Epstein, I. (2002). A comparison of experts' perspectives on assertive community treatment and intensive case management. *Psychiatric Services, 53*, 207–210.
- Scheirer, M.A. (1994). Designing and using process evaluation. In J.S. Wholey, H.P. Hatry, & K.E. Newcomer (Eds.), *Handbook of practical program evaluation* (pp. 48–60). San Francisco: Jossey-Bass.
- Scriven, M. (1997). Empowerment evaluation examined. *Evaluation Practice, 18*, 165–175.
- Teague, G.B., Bond, G.R., & Drake, R.E. (1998). Program fidelity in assertive community treatment: Development and use of a measure. *American Journal of Orthopsychiatry, 68*, 216–232.
- Ziguras, S.J., & Stuart, G.W. (2000). A meta-analysis of the effectiveness of mental health case management over 20 years. *Psychiatric Services, 51*, 1410–1421.

Margaret Gehrs, R.N., M.Sc.N., CPMHN(C), is the clinical leader manager of the Community Mental Health Service of St. Michael's Hospital in Toronto. She also holds a clinical appointment at the Faculty of Nursing, University of Toronto.

Heather Smith Fowler, B.A., M.A., is the research manager of the Women's Health Research Unit at the Institute of Population Health, University of Ottawa.

Sean Rourke, Ph.D., is a scientist with the Centre for Research on Inner City Health and director of research with the Mental Health Service at St. Michael's Hospital, scientific and executive director with the Ontario HIV Treatment Network, and Associate Professor of Psychiatry, University of Toronto.

Donald Wasylenki, M.D., FRCPC, is a professor and chair, Department of Psychiatry, University of Toronto and Psychiatrist in Chief at St. Michael's Hospital, Toronto.

Marnie Smith, B.Sc.N., M.S.W., is a program manager at Canadian Mental Health Association, Ottawa Branch, Ottawa.

J. Bradley Cousins, Ph.D., is a professor of educational administration at the Faculty of Education, University of Ottawa, with a specialization in program evaluation theory and practice. He is the 1999 winner of the Contributions to Evaluation in Canada award of the Canadian Evaluation Society. He has published numerous articles and books in the area of program evaluation. He is currently completing the book, *Participatory Evaluation Up Close: An Integration of Research-Based Knowledge*.

