APHASIOLOGY, 1997, VOL. 11, NO. 4/5, 415–432

Retrospective study of treatment outcome for individuals with aphasia

K. H. ODELL, S. BAIR†, M. FLYNN†,
M. WORKINGER‡, D. OSBORNE§ and M. CHIAL

University of Wisconsin–Madison, WI, USA
† Meriter Hospital, Madison, WI, USA
‡ Marshfield Clinic, Marshfield, WI, USA
§ St Luke's Medical Center, Milwaukee, WI, USA

Abstract

Measurement of outcomes subsequent to treatment and documentation of the efficiency with which outcomes are achieved is critical information for health-care policy makers and third-party payers. This study employed the ASHA Functional Communication Measure (FCM) scales to retrospectively analyse charts of 20 aphasic patients. By discharge, both severe and moderate groups gained a median (across modalities) of 1 FCM level. The severe group remained dependent for communication, while the moderately impaired group typically achieved independent communication levels. Efficiency (amount of FCM level gain relative to number of treatment sessions) was greater for the moderate group; average number of treatment sessions was 40 for the severe group and 22 for the moderate group.

Introduction

Public policy-makers and payers in the external market place make decisions concerning cost, capitation, service delivery, duration of treatment programmes, and proof of value of services (Fratalli 1992). Frequently these decisions are made without adequate supportive data from the pertinent clinical fields. Communicative disorders professionals must inform decision-makers in public and private sectors about the character, costs, and benefits of services provided to persons with disorders and impairments that affect communication.

Of central concern is the issue of treatment outcome, specifically the eventual outcome of the treatment process for the patient, relative to number of treatment sessions. The notion of treatment outcome can be defined as that level of communication ability reached at the point of treatment termination. The term does not imply that this level is the ultimate achievement of the aphasic individual, because communication 'practice', such as that which occurs in daily life, may well result in additional gains in communication ability subsequent to speech–language treatment.

The project reported here, developed under the auspices of the Wisconsin Speech–Language–Hearing Association, examined retrospectively clinical data on communication levels and total treatment sessions collected at a single medical site

Address correspondence to: Katharine H. Odell, PhD, Department of Communicative Disorders, University of Wisconsin–Madison, 1975 Willow Drive, Madison, WI 53705, USA.

0268–7038/97 $12·00 © 1997 Taylor & Francis Ltd

from patients with aphasia. The retrospective nature of the study documents the results of current practice and serves as a base against which to compare outcomes as changes in the health-care delivery system emerge.

Since the publication in 1969 of the Functional Communication Profile (Sarno 1969), various professional groups have developed systems to document or estimate level of ability of aphasic persons to communicate in natural life situations outside the clinic setting. The ongoing development of such systems suggests that clinicians who use them have not been fully satisfied by any one of them.

From among the various measurement devices, the FCM (American Speech–Language–Hearing Association (ASHA 1995)) system was selected for several reasons as the tool to estimate ability to communicate in the real world. The format allows retrospective extraction of information, because it is harmonious with typical assessment procedures and tests, such as the Boston Diagnostic Aphasia Examination (BDAE, Goodglass and Kaplan 1983). Other systems, such as the Functional Independence Measure (FIM) (State University of New York at Buffalo 1993) and the Functional Assessment of Communication Skills for Adults (Frattali *et al.* 1995) demand estimation either on-line with assessment or require clinicians to set up or observe communication scenarios such as telephone conversations. The FCM also describes ability in terms easily understood by lay people, as encouraged by third-party payers who demand descriptions of how well communication abilities can function outside the clinic. Further, as a measure developed by ASHA and beta-phase tested in 150 facilities nationwide, the FCM has good potential for wide distribution and application, thus maximizing opportunities for future cross-study comparisons. In addition, the scales describe a range of component processes of communication (e.g. auditory comprehension) that may be disrupted in aphasia as well as other communication disorders, enabling comparison across disorders. The Appendix reproduces from the ASHA (1995) material the FCM levels and descriptors for the four verbal communication scales.

Although the advantages of using the FCM scales for this study were judged to outweigh the disadvantages, unfavourable features of the FCM are notable, especially the unproved validity and reliability as well as the unknown sensitivity of the scales to subtle improvements in ability. The name of the scale erroneously suggests that it measures 'functional communication', a term now reserved to refer to the success with which non-verbal, non-oral, and verbal abilities are integrated to enable successful communication; although clinicians might be able to estimate functional communication abilities from perusal of the individual subscales of the FCM, the FCM procedures do not specifically facilitate that estimation. The FCM is a preliminary measure, with ongoing development and scrutiny.

This study reports preliminary data on outcome of the treatment process for a group of patients with aphasia. However, attributing outcome change to treatment *per se* is unwarranted, because these subjects received all or part of their treatment during the period associated with spontaneous recovery. Thus, with this caveat in mind, results are discussed in terms of the change in function effected during a period when treatment was administered, but this study is not a documentation of treatment efficacy.

Method

Setting

Charts for the study were selected from among those of discharged patients at a full service community hospital, with an acute inpatient medical ward, an acute inpatient rehabilitation programme (in which patients receive a minimum of 3 hours of treatment a day across disciplines), day treatment programmes, and outpatient service. The clinical staff relevant to this study had professional experience ranging from 3 to 15 years post-award of the Certificate of Clinical Competence with a focus on neurogenics service. The average length of stay for patients was 24 days. Treatment sessions were typically 30 minutes, occurring twice a day, particularly when the patient was transferred to the rehabilitation programme.

Chart selection

For another purpose the medical centre administration had developed a selected list of patients who were diagnosed with a left CVA by choosing every third medical record of such patients in 1994–5. From this truncated list the current study analysed the charts of all patients who met the following inclusionary criteria: diagnosis of aphasia (i.e. received an ICD-9 (US Department of Health and Human Services 1994) code of 784.3); neuroradiographic evidence in the medical chart of a single left cerebral hemisphere CVA; received a full range of service in the facility, from diagnosis/assessment through a treatment phase to discharge; and service limited to 1994–5 to ensure some degree of similarity in third-party payer imposed restrictions on maximum treatment durations. Twenty charts were analysed.

Exclusionary criteria were applied to maximize the homogeneity of the study sample. Charts of 10 patients were rejected for the following reasons: multiple CVAs; co-occurring disorders such as dementia that precluded full and meaningful participation in the treatment process; co-occurring disorders such as apraxia of speech, which received more treatment than the diagnosed aphasia.

Subjects in the severe group were seven females and three males, ranging in age from 48 to 70 years ($\bar{X} = 69$). Subjects in the moderate group were seven females and three males, ranging in age from 60 to 83 years ($\bar{X} = 75$). Additional biographical and medical information is presented in Tables 1 (severe group) and 2 (moderate group).

Chart analysis

Charts were reviewed by a group of three investigators, who are ASHA certified, experienced speech–language clinicians and have used a FIM-like system in clinical practice. Ratings were determined by consensus.

Initial and discharge levels of communication were estimated by application of the appropriate FCM (ASHA 1995) scales to assessment measures and anecdotal speech–language pathologist comments described in the charts. The most common assessment measures employed by these clinicians were the BDAE (Goodglass and Kaplan 1983), the Boston Naming Test (Goodglass *et al.* 1983), and the Token Test

Table 1. Biomedical information on subjects in the severe group

Subject	Concomitant motor speech	Condition			Psychiatric history	Educational history	Discharge status	Post discharge recommendations
		Hemiparesis/ plegia	Sensory loss[a]	Other medical conditions[b]				
S1	AOS?					High school graduate	Daughter's home	Cont. TX
S4		Right		COPD, AODA, bilateral hearing aids, gastrostomy feeding tube		Third grade	Nursing home	Cont. TX
S6	AOS?	Right			History of depression	High school graduate	Home	Cont. TX
S8	No	Right	Right homonymous hemianopsia	Unilateral hearing aid		College+	Home	Cont. TX
S9				CABG, CHF		?	Home	Cont. TX
S10	No	Right	Early right neglect			?	Initially to daughter's home	Cont. TX
S11		Right		Visual impairment; dysphagia; Satzman's syndrome		High school graduate	Home	Cont. TX
S13	No	—		Dysphagia		High school graduate	Home	—
S14	AOS	Right		Dysphagia		College	Home	Cont. TX
S15	No	Right				Some high school	Home	Cont. TX

[a] Consequent to brain damage, such as hemianopsia.
[b] General chronic medical conditions, such as chronic occlusive pulmonary disease.
AOS = apraxia of speech; Cont. TX = continued treatment; AODA = alcohol and other drug abuse; CABG = coronary artery bypass graft; CHF = congestive heart failure; COPD = chronic occlusive pulmonary disease.

Table 2. Biomedical information on subjects in the moderate group

Subject	Concomitant motor speech	Condition: Hemiparesis/plegia	Condition: Sensory loss[a]	Condition: Other medical conditions[b]	Psychiatric history	Educational history	Discharge status	Post discharge recommendations
S2		Right		HTN, CHF, cardiomyopathy			Home	No TX
S3	AOS	Right	—	Cardiac disease			Home	No TX
S5		Right	Right inferior quandranopsia	Cardiac disease, HTN		High school	Nursing home	Cont. TX
S7	No			HTN, COPD, PVD		College	Home	No TX
S12		Right	Right gaze preference	Diabetes, HTN		Tenth grade	Home	Cont. TX
S16	AOS	Right		Cardiac disease, HTN, CHF			Home	Cont. TX
S17	dys	Right		Migraines, HTN, gastric disease		Eighth grade	Home	Cont. TX
S18		Right		Cardiac disease, HTN, CHF, renal insufficiency		High school +	Home	Cont. TX
S19	AOS	Right		HTN, cardiac disease diabetes		High school	Home	Cont. TX
S20		Right		CA	Depression—stable	High school	Home	Cont. TX

[a] Consequent to brain damage, such as hemianopsia.
[b] General chronic medical conditions, such as heart disease.
HTN = hypertension; CHF = congestive heart failure; AOS = apraxia of speech; COPD = chronic occlusive pulmonary disease; PVD = peripheral vascular disease; dys = dysarthria; CA = cancer; TX = treatment; Cont. = continued.

(DeRenzi and Vignolo 1962). The procedure for estimating FCM level involved several steps. Initially, raw data from these measures were scrutinized to reveal subject ability in each modality [auditory comprehension (AC), reading comprehension (RC), oral expression (OE), and written expression (WE)]. Percentile rankings on the BDAE were also identified for subtests in these areas. Chart notes during initial and final assessment periods were perused for mention of abilities. To stabilize the occasionally non-specific language of the FCM performance descriptions at some levels for some modalities, the examiners appended to the FCM language a specific notation of ability at that level, such as 'ability to add a single sentence to a greeting card' or 'ability to compose and write a short note to a friend'. Attention was also given to anecdotal clinical notes on gestural or other augmentative communication ability, as no published or standardized measure was employed for assessment. A typical procedure might unfold as follows: ability to respond accurately to all the paragraph AC questions on the BDAE (Goodglass and Kaplan 1983), plus good skills on the Token Test (DeRenzi and Vignolo 1962), particularly elements of Part V, would be considered AC FCM level 5; however, chart notes reporting difficulty in daily conversations with family might be sufficient evidence to lower the AC FCM to 4.

Usually, the relevant FCM (ASHA 1995) scales were AC, RC, OE, and WE. For only one of the 20 subjects, non-verbal communication was an integral part of the clinical picture at initial and final assessment. Concurrent disorders, such as apraxia of speech, were noted but not rated.

The total number of $\frac{1}{2}$-hour treatment sessions was recorded. In addition, a breakdown of treatment intensity by modality was recorded. The number of treatment sessions each modality received was able to be identified because, fortuitously, the treatment charts included daily notes noting specific goals and tasks. For instance, in a given session, AC, OE and WE might have been targeted; in the 'treatment × modality' summary this session would have been recorded as: 1 AC, 1 OE, and 1 WE, signifying that each modality received treatment once. Many tasks are inherently multi-modality based, such as structured conversation; in these instances the examiners used chart notes to determine the main goal(s) of the task. For conversation, depending on charts notes, this task might be designated for a given subject as either AC or OE or both. Although the precise amount of time per session devoted to treating any modality could not be estimated from the chart notes, the summary figures give an impression of the intensity with which each modality was addressed.

Subjects were grouped according to initial median FCM levels. Levels of 0–2 represent 'severe' impairment; levels of 3–5 represent 'moderate' impairment.[1] This rating process sorted the charts into severe and moderate groups of 10 subjects each.

Individual and group data were collected on four areas of interest:

1. Change from initial to final assessment in overall FCM level (median FCM across the four verbal component processes) and by component process.
2. Description of communication ability at initial and final assessments.
3. Efficiency indices for overall level and by component process. Efficiency indices were developed by dividing change in FCM level by number of treatment

[1] Level 6 represents 'mild' impairment, but no chart with an initial rating that high appeared in either

sessions. Essentially, this index yields a measure of the amount of total change effected during each session.

4. Amount of FCM gain relative to numbers of treatment sessions.

Comparisons between groups and among component processes were analysed by non-parametric statistics, because of the probability for violation of the assumption of normalcy in parametric statistics.

Analysis reliability

The group rated five charts (two severe, three moderate level) a second time, no less than 1 month after the initial rating. Spearman rank-order correlation coefficients (Siegel and Castellan 1988) were calculated for both the median overall initial FCM (I FCM) and the final FCM (F FCM) levels across these charts on Time 1 and Time 2. The Spearman rho for I FCM was 0·94 and for F FCM was 0·91.

Reliability figures for recording the number of times each modality received treatment were determined by tabulating these events a second time for the same five charts that were re-rated (above). The Spearman rank-order correlation coefficients (Siegel and Castellan 1988) between Time and Time 2 tabulations were: AC: $r_s = 0{\cdot}7$; OE: $r_s = 0{\cdot}9$; RC: $r_s = 1{\cdot}0$; WE: $r_s = 1{\cdot}0$.

Results

FCM changes

To identify the amount of FCM level change per group, it was necessary to characterize the overall quality of performance for each subject at initial and final assessment and note the difference between the two. For each subject a median overall (OA) FCM score across the modalities was generated at initial and final assessment. For example, as shown in Table 3, Subject 6 in the severe group received initial ratings of 2, 1, 2, and 2, across the four verbal modalities, yielding an I FCM level of 2. Similar procedures on the final assessment resulted in a F FCM of 3. The change score for this subject, which was 1, was then grouped with that of the other subjects in the severe category, and a median FCM change score for the group was identified. Raw data for the moderate group are detailed in Table 4.

The median change score for both groups was 1. That is, both groups achieved one level of FCM improvement from initial to final assessment, although the ultimate actual level of achievement was higher for the moderate than the severe group.

Description of communication abilities

A frequency distribution of individual subjects' I FCM and F FCM describes more fully than a single-figure median score the level of performance at initial and final assessment. As shown in Figure 1, in the severe group, eight of the 10 subjects had OA median ratings of 2 at initial assessment. That is, across the modalities, communication function was rated at level 2, indicating no independent, spontaneous, consistent communication. The most common rating at final assessment was 3, but one individual reached level 6, and one 5·5.

In the moderate group, the mode (five subjects) at I FCM was 3 and at F FCM was 5 (four subjects). This 3 to 5 difference in modes might suggest that the moderate group made gains of 2 FCM levels, in seeming contrast to the median

Table 3. Functional Communication Measure (FCM) levels at the initiation and termination of treatment for subjects in the severe group. Modalities addressed are: auditory comprehension (AC), oral expression (OE), reading comprehension (RC), written expression (WE), and recognition and expression of non-spoken language (both reception and expression represented as 'non-verbal' NV)

	Component processes of communication								
Subject	AC	OE	RC	WE	NV	Median Overall FCM	FCMΔ	Treatment Sessions	Efficiency
S1									
I FCM	2	2	2	2	—	2			
F FCM	4	4	4	3	—	4	2		
Treatment nos.	28	33	33	27				34	0·06
S4									
I FCM	2	1	0	NA	NA	1			
F FCM	2	2	1	NA	1	2	1		
Treatment nos.	8	10	(1)	(4)				16	0·06
S6									
I FCM	2	1	2	2	2	2			
F FCM	4	3	3	2	4	3	1		
Treatment nos.	42	48	28	23	12			48·25	0·02
S8									
I FCM	3	2	1	2	—	2			
F FCM	4	4	2	2	—	3	1		
Treatment nos.	31	36	11	15	5			33·5	0·03
S9									
I FCM	2	2	1	2	1	2			
F FCM	3	3	3	3	—	3	1		
Treatment nos.	10	14	13	4				18	0·06
S10									
I FCM	2	2	4	2		2			
F FCM	3	4	4	2	—	3·5	1·5		
Treatment nos.	18	23	14	6	(3)			24	0·06
S11									
I FCM	2	1	2	1	—	1·5			
F FCM	3	2	3	2		2·5	1		
Treatment nos.	44	33	20	10	(9)			47	0·02
S13									
I FCM	2	2	2	2		2			
F FCM	6	6	5	5		5·5	3·5		
Treatment nos.	15	58	32	25				59	0·06
S14									
I FCM	3	2	2	2	—	2			
F FCM	6	6	6	4	—	6	4		
Treatment nos.	25	96	23	28	(1)			110	0·04
S15									
I FCM	2	2	3	2		2			
F FCM	3	3	4	2		3	1		
Treatment nos.	9	9	7	2				10·5	0·1

Δ = 'change'; I FCM = FCM level at initial assessment; F FCM = FCM level at final assessment; NA = 'not available'.

Table 4. Functional Communication Measure (FCM) levels for subjects in the moderately impaired group. Modalities addressed are: auditory comprehension (AC), oral expressions (OE), reading comprehension (RC), written expression (WE), and recognition and expression of non-spoken language (both reception and expression represented as 'non-verbal' NV)

	Component processes of communication								
Subject	AC	OE	RC	WE	NV	Median Overall FCM	FCMΔ	Treatment Sessions	Efficiency
S2									
I FCM	4	4	3	2	NA	3·5			
F FCM	6	6	4	2	—	5	1·5		
Treatment nos.	3	19	8	19				25·5	0·06
S3									
I FCM	3	2	3	2	—	2·5			
F FCM	4	3	3	3	—	3	0·5		
Treatment nos.	9	12	9	10				14	0·04
S5									
I FCM	3	3	4	2	—	3			
F FCM	4	4	4	2	—	4	1		
Treatment nos.	11	11·25	2	1				12	0·08
S7									
I FCM	4	3	3	3	—	3			
F FCM	6	6	6	6	—	6	3		
Treatment nos.	11	21·5	13	10				25	0·12
S12									
I FCM	3	4	3	2	—	3			
F FCM	5	6	4	3	—	4·5	1·5		
Treatment nos.	3	15	18	9				18	0·08
S16									
I FCM	3	3	3	2	—	3			
F FCM	4	5	4	3	—	4	1		
Treatment nos.	5	15	6	10	(3)			28	0·04
S17									
I FCM	4	4	4	3	—	4			
F FCM	5	5	5	4	—	5	1		
Treatment nos.	9	8	4	8				16	0·06
S18									
I FCM	3	3	3	3	—	3			
F FCM	6	6	5	5	—	5·5	2·5		
Treatment nos.	19	37	29	13				33·5	0·07
S19									
I FCM	4	2	4	NA	—	4			
F FCM	5	5	5	3	—	5	1		
Treatment nos.	5	40	18	1				41	0·02
S20									
I FCM	3	4	4	5	—	4			
F FCM	5	6	5	5	—	5	1		
Treatment nos.	4	9	5	5				15	0·07

Δ = 'change'; I FCM = FCM level at initial assessment; F FCM = FCM level at final assessment; NA = 'not available'.

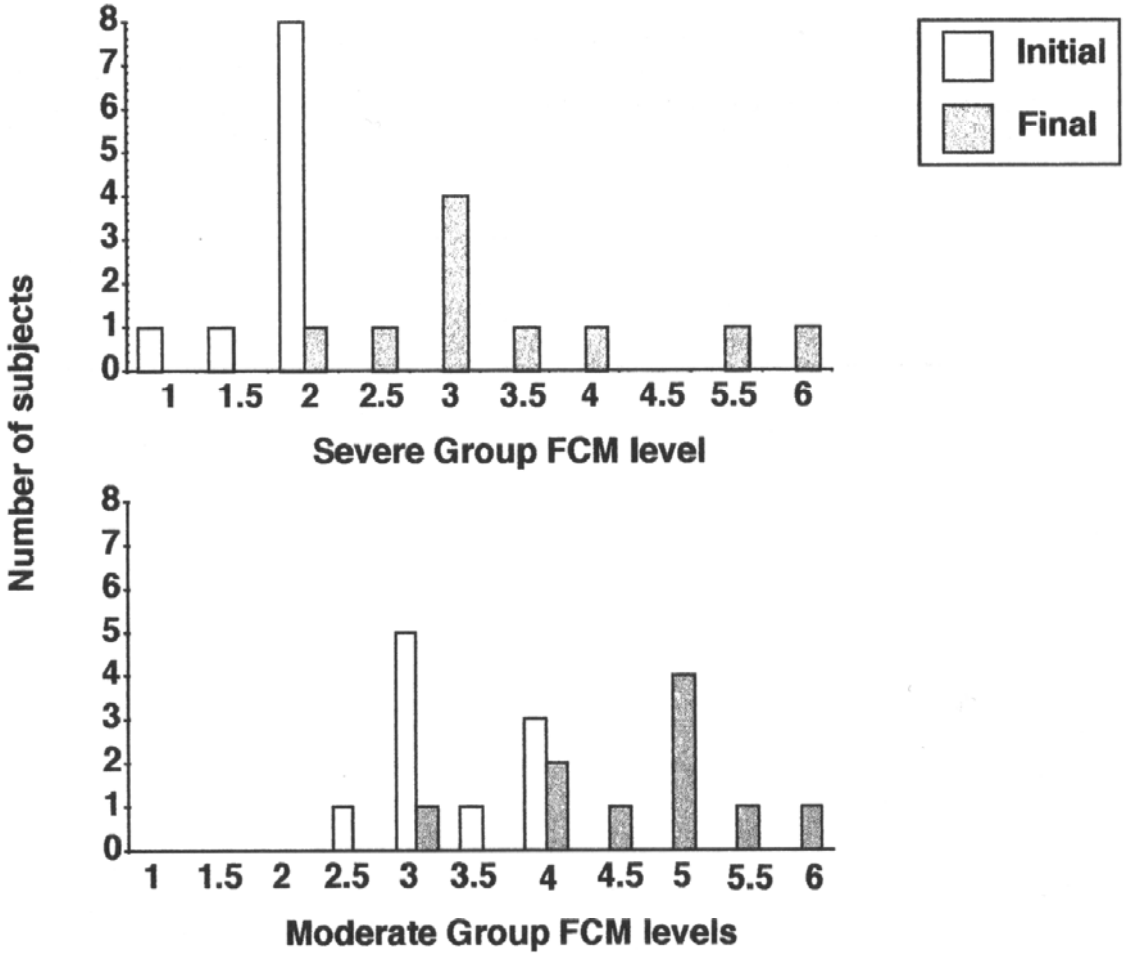

Figure 1. Frequency distribution of initial and final FCM levels for the two groups.

gain figure noted above. However, the frequency distribution points out better than a median score that four subjects were at levels below 5 at final assessment; these lower ratings illustrate how group median change score of 1, as noted above, is the most suitable single score indication of moderate group gains but does not present the full picture.

Outcome at discharge left subjects in each group with quite different communication abilities. Ability at level 3, the level of final achievement for most severe subjects, reflects limited spontaneous, independent communication; most communication occurs at the word and phrase level and requires repair or assistance from communicative partners. Level 4, at which 20% of the moderate group were left, denotes sentence-level ability but continuing limitations in complexity of content or language code use. Level 5, describing 40% of the moderate subjects, reflects notably greater independence in communication, more spontaneity and management at the sentence-level, but still restricted in variety and communication efficiency.

Median FCM gains within each verbal modality (AC, OE, RC, WE) were derived in the following fashion: finding the median among all initial AC ratings within a group, then finding the median among all final AC levels, and subtracting the initial from the final score. By this calculation, in the severe group, median gains were equivalent, at 1·5 levels, for all modalities except WE, which made no median gain. In the moderate group, most gains were noted in OE (median = 2·5

Table 5. Tabulation and percentage of sessions according to whether one, two, or more modalities were addressed in each session. Tabulation represents data from five subjects. Only the primary language modalities AC (auditory comprehension), OE (oral expression), RC (reading comprehension) and WE (written expression) are reported

	Modalities											
	AC			OE			RC			WE		
Subject	1	2	3	1	2	3	1	2	3	1	2	3
Severe												
S13	1	13	2	11	46	2	1	29	2	0	21	2
S14	0	24	0	44	49	0	0	30	0	1	28	0
Totals[a]	1	37	2	55	95	2	1	59	2	1	49	2
Summary %[b]	/	93%	/	36%	63%	/	/	95%	/	/	94%	/
Moderate												
S5	1	15	1	7	18	1	0	2	1	0	0	1
S16	0	6	0	1	13	0	0	6	0	0	10	0
S17	0	9	0	0	4	0	0	4	0	0	8	0
Totals	1	35	1	8	35	1	0	12	1	0	18	1
Summary %	/	95%	/	18%	80%	/	0	92%	/	0	95%	/

[a] Totals = Total number of sessions across subjects within that level (e.g. single modality, multiple modalities).
[b] Summary % = percentage of the total number of sessions represented by sessions in that level.

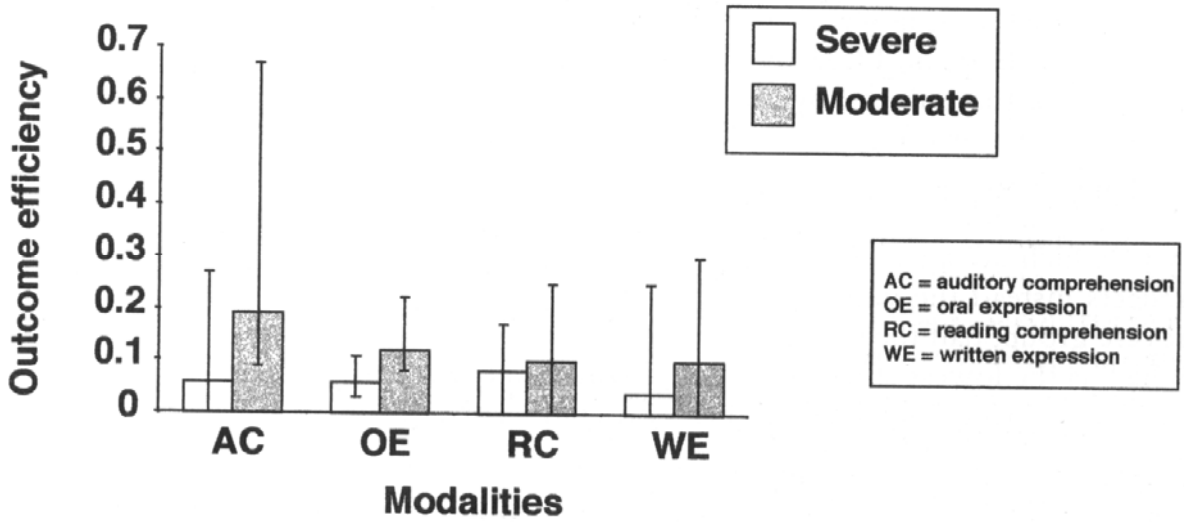

Figure 2. Median efficiency indices and ranges across the four verbal modalities for the two groups.

levels), then AC (median = 2 levels). These figures indicate that the median overall gain of 1 level noted in both groups was not restricted to one modality.

Efficiency of outcome gains

Efficiency, or the rapidity with which the gains were accomplished, was calculated by weighting the median FCM level change per subject by the average number of treatment sessions for that group, and ultimately deriving a group figure. Average number of treatment sessions was 40 for the severe group and 22 for the moderate group. Larger efficiency indices mean greater efficiency, a greater amount of gain

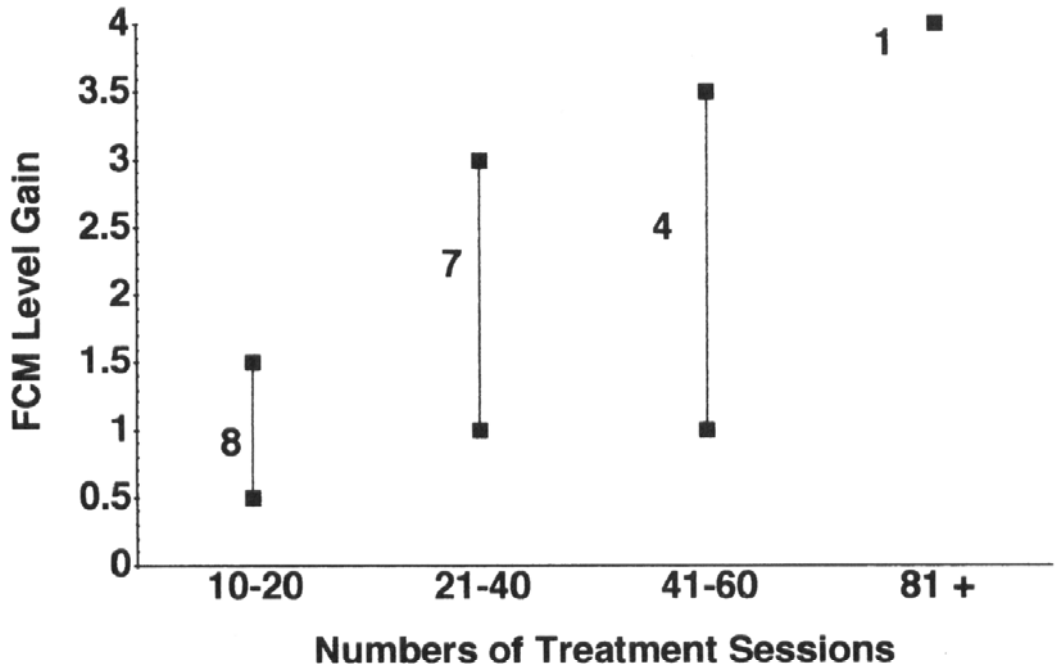

Figure 3. Ranges of FCM gain as a function of aggregate number of treatment sessions. Numbers beside each range bar indicate number of subjects in that category.

per session. The efficiency indices for the groups, 0·03 for the severe and 0·05 for the moderate group, were statistically different, according to a Wilcoxon–Mann–Whitney test (Siegel and Castellan 1988), corrected for ties. This result is consistent with the expectation that treatment is more efficient for the moderate subjects.

Other methods, however, exist to examine individual and group efficiency. One entails dividing the FCM overall gain per subject by the number of treatment sessions for that subject. By this calculation the median of the ranges in each group (0·060 for the severe group and 0·065 for the moderate) do not differ statistically, by a Wilcoxon–Mann–Whitney test (Siegel and Castellan 1988), corrected for ties. This calculation does not well represent differences between groups in the speed with which advances were made; this median seems to obscure differences and reflects instead the wide range of efficiency scores. As shown in Table 1, in the severe group, half the subjects received scores of 0·06; two received scores of 0·02, which is the lowest efficiency score obtained in either group. More subjects in the moderate group obtained higher efficiency scores—one at 1·2 (3 levels in 25 sessions), as detailed in Table 2.

To discover which modalities benefited most from the treatment period, efficiency scores per modality were determined. For each subject, amount of FCM level gain in AC, for instance, was weighted by the number of treatment sessions devoted to it for that subject. Once AC efficiency scores for all subjects in each group were developed, a median for that group was derived. As shown in Figure 2, the groups differed in which modalities showed the greatest gain; AC in the moderate group, but RC in the severe group. Consistent with the amount of gain per modality analysis, WE for both groups was either last or tied for last place in this comparison.

FCM gain as a function of aggregated treatment sessions

The last analysis looked at the relationship of the amount of change in FCM level relative to aggregated treatment categories. As shown in Figure 3, subjects

participating in more treatment sessions did gain more FCM levels, regardless of initial FCM level. Aggregation of treatment sessions into the four categories shown was based primarily on average numbers of treatment sessions for the two groups (22 for the moderate, 40 for the severe). Those who received one to 22 sessions gained 0·5–1·5 levels; this segment of the sample included three severe and five moderate subjects. Three severe and four moderate subjects receiving 23–40 sessions gained 1–3 levels. Four subjects (three severe, one moderate) treated for 41–60 sessions gained a similar amount to those receiving 22–40 sessions. The one fortunate subject (in the severe group) receiving 110 sessions gained the most, 4 levels, enabling him to return to his previous employment as an architect.

Discussion

This study provided information on the degree to which communication in a group of subjects with aphasia was effective at final assessment post-treatment. Third-party payers now demand documentation that treatment produces meaningful changes in an individual's communicative life; this new orientation requires either procedures different from those earlier ones relying on standardized assessment, or different techniques to interpret aphasic performance from these standardized tests. While standardized tests, of which the BDAE is only one example, were not designed to facilitate the translation from assessment measure to level of communication effectiveness, it appeared feasible, as well as possibly valid and reliable, to interpret retrospectively the level of communicative effectiveness from these measures. Although prospective studies of communication effectiveness are in operation, the urgent need to clarify the value of speech–language services is apparent. Thus, this study undertook a retrospective analysis of patient charts to obtain some preliminary understanding of the speed and outcomes of treatment, described in terms more accessible to consumers and third-party payers.

One benefit of the retrospective nature of the study is the analysis of a natural clinical situation; no limits or expectations were placed on treatment durations or intensity, permitting exploration of aspects of the treatment period as they occurred naturally. One disadvantage of naturalistic studies is lack of control over the criteria for discharge. In this pool of subjects, most were discharged from the medical centre's speech pathology service based on decisions made by the family or other health-care disciplines. Speech–language pathologists typically recommended further treatment, but the subjects were terminated from the medical centre programme for various reasons (e.g., to return home). With further therapy, outcomes might well have increased.

The FCM scale was found to be adequate to note changes in the subjects studied. The four verbal FCM scales were appropriate for the study of the disability in language use in aphasia. Other disorders, however, such as the cognitive–communicative disabilities of subjects with right brain damage or the speech disabilities of those with laryngectomy, might not be as well served by the FCM.

One concern voiced by clinicians regarding many of the communication rating measures, potentially including the FCM, is that a seven-interval scale does not capture subtle gains that signal improvement; small advances may be clinically important, but be too limited to advance the subject to the next level on the scale. This study found few instances when gains were noted in the charts (nine out of a possible 81 instances to note change) but could not be registered as a move up on

the scales. For example, for moderate Subject 3, no change was registered for RC, despite improvement from 40% to 60% in sentence level comprehension. Similarly, severe Subject 6 advanced in WE from writing only her own name spontaneously to writing independently names of family members; this improvement rated no FCM change. In only one subject (severe S4) was no change noted in performance in the chart. The improvements noted in the eight cases were not only too weak to be noted on the FCM but also judged insufficient to argue for continued treatment, if that course had been considered. The examiners felt that these scales and the descriptions of performance at each level not only corresponded to performance abilities actually displayed by patients, but also captured adequately the kinds of gain typically demonstrated.

Nevertheless, weaknesses of the FCM at this time dilute full confidence in its ability to register subtle clinical changes. Despite apparent face validity the scale has not received sufficient analysis of validity. Furthermore, the scale is an ordinal one, which may include some levels that encompass a wide range of behaviours, thus obscuring some noted changes, while other levels include a more narrow range, that magnifies other subtle changes. Recently, investigators (Linacre *et al.* 1994) have applied RASCH analysis to the FIM (State University of New York 1990), to statistically transform the finite and restricted range of that ordinal scale into a more expansive range of interval scale; this technique may be of value with the FCM to refine its ability to register at each successive level equal amounts of improvement, or additions, in communication ability.

Travel in the frontier is not always clear and unambiguous, and the current investigative process is no exception. Regarding data analysis, many different sets of rating numbers can be corralled to derive median scores to reflect typical performance. For instance, the method used in this study to note amount of change per group was to secure the median of all the change scores per group; however, another technique is to obtain a median of all the I FCMs, then a median of all the F FCMs, and subtract these medians. Because different sets of numbers are used, these two final medians are different for the moderate group, resulting in a conundrum for data analysis.

These data showed that severe subjects who started at FCM levels of 1 or 2 typically gained 1 FCM level but required a treatment period of about 40 $\frac{1}{2}$-hour sessions. Similar amounts of improvement occurred in the moderate group, but more rapidly, associated with fewer treatment sessions. Starting at lower levels of ability did not negatively affect the amount of gain achieved in these subjects.

This amount of gain for moderate subjects, although not dramatic by itself, was heartening because, at final assessment, communication was functional and relatively independent. Although severely impaired subjects did gain—and two gained a remarkable amount—the ultimate improvement experienced still left most at moderate levels of communication, requiring assistive cueing to address primary daily needs. The burden of communicating with these individuals remains on their communicative partners; with greater dependency there is less likelihood of successful reintegration into the community and of achieving healthy personal well being (Lyon 1992).

These data show that more treatment results in more gain. This situation is positive, although frustrating if coverage for longer treatment periods cannot be secured. Speech–language pathologists can legitimately strive to minimize the chance that they will treat only to the point where they can document one FCM

level change. Passive acceptance of a cap at 22 sessions—or less—on the number of sessions is accepting the lowest denominator. However, the salad days of the past are over—treatment periods are and will be shorter than in 1985 or in 1995. Thus, to maximize recovery, it behooves clinicians to design methods to extend treatment in cost-saving ways. Obvious techniques include group programmes or activities executed by the patient at home, with assistance from family and friends and monitoring by speech–language clinicians. Group and home-based programmes are currently common clinical tools, however, current documentation is lacking regarding individual programme efficacy or efficiency of these efforts relative to communication outside the clinical environment, notwithstanding the extensive work done by Wertz and colleagues in the 1980's on treatment efficacy in various settings (Wertz *et al.* 1981, 1986).

Determination of efficiency of treatment outcome must be based on both time (number of treatment sessions) as well as amount of gain and ultimate achievement of functional communicative level. Although a singular focus on time—i.e. the 'time as money' issue—is the present perspective, this view ignores the critical aspect of treatment, which is to foster independence in communication, as can be indicated by gains in FCM level. The health-care policy pendulum may equilibrate somewhat within next half-decade to reducing the emphasis on cost, and taking a more considered look at quality for cost. More data will add to a comprehensive picture of, at least, treatment outcome but ideally also of treatment efficacy and efficiency. This study continues data collection at other Wisconsin sites, examining practices potentially influenced by the span of urban to rural sites and third-party payer territories.

References

American Speech–Language–Hearing Association (1995) *Functional Communication Measure* (ASHA, Rockville, MD).

Department of Health and Human Services (1994) *The International Classification of Diseases*, 9th revision, Clinical Modification (ICD-9-CM). DHHS Publication No. PHS 80-1260 (US Government Printing Office, Washington, DC).

DeRenzi, E. and Vignolo, L. (1962) The Token Test: A sensitive test to detect receptive disturbance in aphasics. *Brain*, **85**, 665–678.

Frattali, C. (1992) Functional assessment of communication: merging public policy with clinical views. *Aphasiology*, **6**, 63–83.

Frattali, C., Thompson, C., Holland, A., Wohl, C. and Ferketic, M. (1995) *Functional Assessment of Communication Skills for Adults*. (American Speech–Language–Hearing Association, Rockville, MD).

Goodglass, H. and Kaplan, E. (1983) *The Assessment of Aphasia and Related Disorders*, (2nd edn). (Lea & Febiger, Philadelphia, PA).

Goodglass, H., Kaplan, E. and Weintraub, S. (1983) *The Boston Naming Test* (Lea & Febiger, Philadelphia, PA).

Linacre, J., Heineman, A., Wright, B., Granger, C. and Hamilton, B. (1994) The structure and stability of the Functional Independence Measure. *Archives of Physical Medicine and Rehabilitation*, **75**, 127–132.

Lyon, J. (1992) Communication use and participation in life for adults with aphasia: the scope of the problem. *American Journal of Speech–Language Pathology*, **1**, 7–14.

Sarno, M. (1969) *Functional Communication Profile*. (New York University Medical Centre, New York).

Siegel, S. and Castellan, N. J. (1988) *Nonparametric Statistics for the Behavioral Sciences*, 2nd edn (McGraw-Hill, New York).

STATE UNIVERSITY OF NEW YORK AT BUFFALO (1990) *Guide for Use of the Uniform Data Set for Medical Rehabilitation: Functional Independence Measure* (State University of New York at Buffalo, Research Foundation, Buffalo, NY).

WERTZ, T., COLLINS, M., WEISS, D., KURTZKE, J., FRIDEN, T, BROOKSHIRE, R., PIERCE, J., HOLTZAPPLE, P., HUBBARD, D., PORCH, B., WEST, J., DAVIS, L., MATOVTICH, V., MORLEY, G. and RESSUCCION, E. (1981) Veterans Administration cooperative study on aphasia: a comparison of individual and group treatment. *Journal of Speech and Hearing Research*, **24**, 580–594.

WERTZ, T., WEISS, D., ATEN, J., BROOKSHIRE, R., GARCIA-BUNUEL, L., HOLLAND, A., KURTZKE, J., LAPOINTE, L., MILITANI, F., BRANNEGAN, R., GREENBAUM, H., MARSHALL, R., VOGEL, D., CARTER, J., BARNES, N. and GOODMAN, R. (1986) Comparison of clinic, home, and deferred language treatment for aphasia. *Archives of Neurology*, **43**, 653–658.

Appendix: Functional Communication Measure (ASHA Task Force on Treatment Outcome and Cost Effectiveness 1995)

Level	Comprehension: spoken language	Production: spoken language	Comprehension: written language	Production: written language
0	Unable to test.	Unable to test.	Unable to test.	Unable to test.
1	No comprehension of spoken language.	No meaningful spoken language.	No comprehension of written language.	No production of written language.
2	Comprehension of spoken language is limited to familiar words and/or phrases related to personal needs, although most responses are inaccurate or inappropriate.	Spoken language is limited to automatic and/or imitative words and phrases, although most attempts are inaccurate or inappropriate.	Comprehension of written language is limited to familiar words and/or phrases related to personal needs, although most responses are inaccurate.	Production of written language is limited to the copying of numbers, letters, and/or words, although most attempts are distorted; spontaneous productions are inaccurate and/or illegible.
3	Comprehension of spoken language consists primarily of simple statements about personal topics, although repetition and/or rephrasing are required. Accuracy of comprehension is erratic.	Spoken language consists primarily of automatic speech with inconsistent words or phrases, although production may be accurate in imitation.	Comprehension of written language consists primarily of words about personal topics, although cueing is required. Accuracy of comprehension is erratic.	Spontaneous production of written language consists primarily of overlearned and/or familiar words, although copying may be normal.
4	Comprehension of spoken language is limited to the primary activities of daily living needs and simple ideas and frequently requires repetition and/or rephrasing.	Spoken language is limited to the communication of primary activities of daily living needs and simple ideas.	Comprehension of written language is limited to the primary activities of daily living needs and simple ideas; frequently requires cueing.	Production of written language is limited to words for activities of daily living and simple ideas.

Appendix: (*cont.*)

Level	Comprehension: spoken language	Production: spoken language	Comprehension: written language	Production: written language
5	Comprehension of spoken language is normal for activities of daily living, but limited in complexity of form, content, or use; self-monitoring is inconsistent.	Spoken language is functional for activities of daily living, but limited in complexity of form, content, or use; self-monitoring is inconsistent.	Comprehension of written language is functional for activities of daily living, but limited in complexity of form, content, or use; self-monitoring is inconsistent.	Production of written language is functional for activities of daily living, but limited in complexity of form, content, or use; self-monitoring is consistent.
6	Comprehension of spoken language is normal in most situations, although minimal difficulty may occur; self-monitoring/self-correcting are present.	Spoken language is functional in most situations, although minimal difficulty may occur, e.g., word recall, latency of responding; self-monitoring/self-correcting are present.	Comprehension of written language is functional in most situations; minimal difficulty may occur; self-monitoring/self-correcting are present.	Production of written language is functional in most situations, although minimal difficulty may occur; self-monitoring/self-correcting are present.
7	Comprehension of spoken language is normal in all situations.	Spoken language is normal in all situations.	Comprehension of written language is normal in all situations.	Production of written language is normal for all situations.