THE ROTER INTERACTION ANALYSIS SYSTEM (RIAS) IN ONCOLOGICAL CONSULTATIONS: PSYCHOMETRIC PROPERTIES

L.M.L. ONG[a,*], M.R.M. VISSER[a], I.P.M. KRUYVER[b], J.M. BENSING[b], A. VAN DEN BRINK-MUINEN[b], J.M.L. STOUTHARD[c], F.B. LAMMES[d] and J.C.J.M. DE HAES[a]

[a] *Academic Medical Center*, *Department of Medical Psychology*, *Amsterdam*, *The Netherlands*
[b] *Netherlands Institute of Primary Health Care*, *Utrecht*, *The Netherlands*
[c] *Merwede Hospital*, *Department of Internal Medicine*, *Dordrecht*, *The Netherlands*
[d] *Academic Medical Center*, *Department of Obstetrics/Gynaecology*, *Amsterdam*, *The Netherlands*

SUMMARY

One of the most frequently used systems to analyse doctor–patient communication is the Roter Interaction Analysis System (RIAS). However, it has mostly been applied and evaluated in primary care settings. Two studies are presented in which the psychometric properties of the RIAS are investigated in an oncological setting.

In the first study ($N = 25$) the feasibility, inter-rater reliability and content validity of the RIAS was investigated. In the second study, we evaluated the discriminant validity of the RIAS.

Results of the first study showed that coding of tapes was more time consuming than indicated by the Roter manual. The inter-rater reliability proved to be high for both physician communication (0.68–1) and patient communication (0.60–1). The content validity proved to be acceptable: all utterances could be classified. However, coding dilemmas regarding affective communication occurred. The RIAS provided no option to classify communication of a third person present. Some communication categories were never or rarely used.

Results of the second study indicate that the RIAS was able to discriminate between communicative behaviors in oncological consultations ($N = 60$) and three different GP-samples (random-GP sample $N = 329$, hypertension sample $N = 103$, gynaecology sample $N = 65$). To conclude, the psychometric properties of the RIAS are satisfactory in an oncological setting. © 1998 John Wiley & Sons, Ltd.

INTRODUCTION

During the past decades, communication between doctors and patients has attracted an increasing amount of attention within health care studies. Descriptive as well as experimental studies have tried to shed light on the communication process during medical encounters. Results from some of these studies indicate that certain aspects of medical communication may have an influence on patients' satisfaction, adherence to treatment, recall and understanding of medical information, coping with the disease, quality of life, and even state of health (Carter *et al*., 1982; Molleman *et al*., 1984; Fallowfield *et al*., 1986, 1990; Kaplan *et al*., 1989; Bensing, 1991a,b; Siminoff and Fetting, 1991; Hall *et al*., 1994a,b). Nevertheless, the insight gained from these efforts is still limited (Ong *et al*., 1995). Studying the interactive behaviors of doctors and patients is essential in order to enhance the outcomes of patient care.

In order to study communication between doctors and patients, social scientists have developed observation instruments, called 'interaction analysis systems'. These systems enable the methodic identification, categorization, and quantification of salient features of doctor–patient communications (Wasserman and Inui, 1983). A differentiation can be made between 'cure' and 'care' oriented systems. 'Cure' oriented systems are meant to capture the instrumental (task focused) communication. 'Care' systems, on the other hand, are meant to measure the more affective (socio-emotional) communication (Bensing, 1991a,b). Besides the cure–care distinction, interaction analysis systems differ from each other

---

* Correspondence to: Academic Medical Center, Department Medical Psychology, J4-408, Meibergdreef 15, 1105 AZ Amsterdam-Z.O, The Netherlands. Tel: + 31 20 5664661; Fax: + 31 20 5664440; E-mail: L.m.ong@amc.uva.nl

with regard to their setting specificity (designed for analyzing communication between doctors and hospitalized patients, in an outpatient clinic or a general practitioners setting), the observational strategy (coding from videotape, audiotape, direct observation or literal transcripts), and the channels of communicative behavior; applicable to verbal, non-verbal behavior or both (Wasserman and Inui, 1983; Ong *et al.*, 1995).

Several criteria play a role when selecting a system for the analysis of doctor–patient communication. Firstly, the interaction analysis system should capture both cure and care oriented behaviors, since communication researchers agree about the importance of both in medical communication (Bensing, 1991a,b). Secondly, an interaction analysis system designed for coding directly from audio or videotape without using literal transcripts or an electronic device is preferred since it allows coding of nonverbal communication, such as intonation. Thirdly, easy to employ recorders are preferred for practical reasons. Regular use of video cameras may be intrusive, especially in an emotionally threatening situation such as an oncological setting. Fourthly, the coding of consultations should not be too time consuming. Fifthly, the interaction analysis system must have been tested for reliability and validity.

The Roter Interaction Analysis System (RIAS) meets these criteria and is one of the most commonly used systems for analysing doctor–patient communication (Roter *et al.*, 1988). It is well documented (Roter, 1991), it proved to be reliable in the USA (Roter *et al.*, 1991; Hall *et al.*, 1994a,b), the UK (Ford *et al.*, 1995) and in The Netherlands (Bensing, 1991a,b), and it was judged favourably in a comparison study (Inui *et al.*, 1982). More importantly, previous studies applying the RIAS have yielded data about the effect of instrumental and affective behaviour on the quality of care and patient satisfaction (Bensing, 1991a,b). The RIAS however, is less suitable for determining the quality of doctor–patient communication for individual consultations. The system is not originally developed to distinguish between a 'good' and a 'bad' consultation. Rather, the data derived from studies using the RIAS could contribute to the development of a theory regarding doctor–patient communication.

The RIAS is derived from Bales' Interaction Process System. It consists of a detailed classification system which is meant to assign each recorded verbal utterance of both doctor and patient to a category. A verbal utterance is defined as the smallest meaningful and distinguishable speech segment, conveying only one thought or relating to one item of interest. This can be one single word or a lengthy sentence (Roter, 1991; Hall *et al.*, 1994a,b). Thirty-four categories are distinguished to classify physician communication and 28 to classify patient communication (Roter manual, 1991). Direct coding makes assessment of the tonal qualities (voice tone, intonation) and emotional context of the visit possible. Besides the coding of utterances, both doctor and patient are rated on five global affect scales. These 6-point scales are meant to assess the following types of affect: anger/irritation; anxiety/nervousness; dominance/assertiveness; interest/engagement; friendliness/warmth (Roter, 1991).

In order to facilitate interpretation of the data, the behavioural categories were grouped into clusters (Hall *et al.*, 1987; Roter *et al.*, 1987). The four instrumental clusters were: giving directions; asking questions; giving information; and, counselling. Further analyses of the affective cluster by Bensing (1991a,b) produced four socio-emotional factors, namely: social behaviour; verbal attentiveness; showing concern; and negative talk.

Most studies using the RIAS examined the content of doctor–patient interactions in a primary care, ambulatory care or general practitioners setting (Hall *et al.*, 1987, 1994a,b; Roter *et al.*, 1987; Bensing, 1991a,b). Only recently, researchers have started to study communication between patients and medical specialists. It can, however, be expected that the content of these consultations is different from doctor–patient communication in a general practice setting. Cancer consultations, for instance, can be assumed to be more emotional. It concerns a life threatening situation where complex decisions have to be made. Concerns, worries and empathic statements are probably expressed to a greater extent by both oncologists and patients. In consultance with the author of the RIAS, it has therefore been decided to add a sixth scale, 'signs of distress', to the global affective scales for both patients and physicians.

In a study by Waitzkin (1985), it was found that the length of acquaintance between doctors and patients was positively associated with doctors' information-giving behaviours. According to this finding, one could expect general practitioners

to be more informative. In the same study however, results indicated that patients with unfavourable—although not necessarily fatal—prognoses tended to receive more information from their specialists (Waitzkin, 1985). Results from various studies have indeed shown that cancer consultations are highly informative (Chaitchik et al., 1992; Eussen et al., 1992; Castejón et al., 1993; Fallowfield et al., 1995; Ford et al., 1995). Oncologists are confronted with a type of illness which requires presentation of detailed information, besides their expression of concern and empathy.

Furthermore, since cancer patients' need for information is high (Blanchard et al., 1988; Fallowfield et al., 1995), it can be assumed that cancer patients' information-seeking behaviours, in terms of question-asking (Sutherland et al., 1989), are higher than general practice patients.

Finally, the cancer consultation is probably less characterized by physicians' intentions to persuade, influence or change patients' behaviours. General practitioners are more often confronted with chronic illnesses, such as high blood pressure or diabetes. Treatment of these illnesses often requires certain alterations in patients' lifestyles. Consequently, GP's probably express more directive communication.

One recent study in the UK evaluated the intercoder reliability, the structure and content of the 'bad news' cancer consultation by applying the RIAS (Ford et al., 1995). However, apart from the reliability, the psychometric properties of the RIAS have not been evaluated in an oncological setting.

We will therefore report two studies. The first study aims at investigating the RIAS: (a) feasibility; (b) inter-rater reliability; and, (c) content validity.

In the second study we will evaluate the discriminant validity of the RIAS: communication during oncological consultations will be compared with communication between patients and general practitioners. The general practice consultations were collected earlier by researchers from the Netherlands Institute of Primary Health Care (NIVEL). Their data were made available for the purpose of evaluating the discriminant validity of the RIAS.

On the basis of previous research and our own experience, it is plausible to expect that oncologists: give more information; show less counselling (statements intending to persuade or influence the patient's behaviour); are more verbally attentive (expressing agreement or empathy, paraphrasing); and express more concern (worry or optimism) as compared to general practitioners.

Cancer patients are expected to: ask more questions and express more concern than general practice patients.

Regarding the global affective scales, it is expected that oncologists have a higher rating regarding the 'interest/engagement' and 'friendliness/warmth' scales as compared to general practitioners. Cancer patients are expected to have a higher rating regarding the 'anxiety/nervousness' scale as compared to general practice patients.

STUDY 1: FEASIBILITY, INTER-RATER RELIABILITY AND CONTENT VALIDITY OF THE RIAS

Method

Sample. Five gynaecologists (one woman, four men) and one (male) urologist recorded initial consultations with cancer patients. Twenty-five consecutive patients (22 women, 3 men) were included. The mean age was 55 years (range 28–83 years). The gynaecological cancers concerned malignancy of the cervix, vulva, uterus, endometrium and ovary. The urological cancers were all prostate malignancies.

During the recorded consultations, the following topics were usually discussed: reason for referral to a specialist; the specific aspects of the diagnosis; medical history; findings of the physical examination; the proposed treatment policy—which is usually surgery; prognosis; side effects and other consequences of the treatment; consequences for bladder and/or bowel function; sexuality; fertility; and, work. The physical examination which took place during these consultations was not recorded for technical and privacy reasons.

Instrument. The RIAS was used to analyze the taped oncological consultations. A sixth scale, 'signs of distress', has been added to the global affective scales for both patients and physicians.

Procedure and analyses. After the physicians informed their patients about the study, they asked for their consent. When patients gave their consent, the physicians recorded their consultations using two small Sony tape recorders: one tape for the patient to take home, one tape for the researcher.

The twenty-five consultations were coded independently by two raters. These raters received the RIAS training by the original author (D.R.) of this system. This one-week training consisted of: an introduction about the development of the RIAS; getting acquainted with the RIAS manual which gives a detailed description of all communication categories and 'rules of thumb'; getting acquainted with the coding forms. Then, six American audiotaped consultations—together with their complete transcripts and correct codes—were used for training purposes. The original manual (Roter, 1991) was used for analyzing our oncological consultations.

Since the RIAS was not developed originally to classify communication of three speakers, only contributions of physicians and patients was analyzed. 'Simultaneous communication' was classified separately for each speaker.

To assess the feasibility of the RIAS, the raters recorded both the length and the coding time of the taped consultations.

For further statistical analyses, sumscores were calculated by adding all utterances (in the form of tallies) within each behavioural category. Firstly, a two-sample *t*-test was performed to determine whether there was a difference between the mean number of utterances for both raters. Secondly, difference scores were calculated between both raters for all categories and clusters. Thirdly, Spearman Correlation Coefficients between both raters were calculated per category and per cluster.

To evaluate the content validity of the RIAS, it was determined whether all utterances could be classified and if coding conflicts occurred. Finally, it was registered if there were categories which were never or rarely used.

Results

The feasibility of the RIAS: coding time. The RIAS manual states that the 'average time' for coding an encounter is usually double the duration of the consultation. In our study, the mean duration of the recorded consultations was 10.5 min (S.D. = 3.3). The average coding time was 47.2 min for rater 1, and 52.4 min for rater 2. In other words, coding time was about five times the duration of the consultation.

It was also observed by the raters that the presence of a spouse or other family member during the consultation influenced the feasibility of coding. Communication between speakers sometimes took place simultaneously. This made coding more complex since the communication was less audible. Especially when the doctor and the patient were of the same sex, their voices became more difficult to distinguish.

The inter-rater reliability of the RIAS. The mean total number of utterances per consultation was nearly identical for both raters, namely 219.2 (S.D. = 113.3) for rater 1, and 216.8 (S.D. = 114.2) for rater 2 ($t(48) = 0.07$, $p = 0.94$). The median of the difference scores for the total number of utterances for both raters was -2 (range: -21 to 10).

For eight RIAS categories and one global affective scale, the raters did not differ in their ratings. These were: physician gives compliments; physician expresses disapproval; physician gives reassurance; patient gives personal remarks/social conversation; patient gives direction/orientation; patient asks for repetition; patient asks for information other than medical/therapeutical; patient asks for services; and, signs of distress by physician (global affective scale).

The eight RIAS categories which showed the largest range of difference scores between rater 1 and rater 2 were: physician transitions (median: -1, range: -22 to 2); physician gives information about lifestyle and psycho-social feelings (median: 0, range: -1 to 18); showing agreement and understanding by the patient (median: 0, range: -4 to 12); patient transitions (median: 0, range: -7 to 2); physician gives therapeutical information (median: -1, range: -6 to 2); unintelligible communication by patient (median: 0, range: -4 to 4); physician gives medical information (median: 0, range: -4 to 3); and, physician shows agreement and understanding (median: 0, range: -2 to 4).

For physician-communication, the inter-rater correlations ranged from 0.68 to 1.00 for the RIAS categories, and from 0.86 to 0.99 for the clusters (Table 1). Concerning patient-communication, the inter-rater correlations ranged from

0.60 to 1.00 for the RIAS categories, and from 0.76 to 0.99 for the clusters (Table 1).

For the global affective rating scales, the inter-rater correlations ranged from 0.55 to 1.00 for the physician ratings, and 0.75 to 0.86 for the patient ratings (Table 2).

The content validity of the RIAS: categories missing or never/rarely used. All utterances of both oncologists and cancer patients during the consultation could be classified in one of the RIAS categories; these categories covered the contents of the consultations adequately.

Classification for some utterances was however questionable. Coding conflicts occurred where two different categories seemed appropriate. This concerned mainly affective utterances expressing concern or optimism. Statements of doctors and patients frequently contained medical information as well as affective communication. For example 'Am I going to die?' could either be coded as 'patient asks medical question' or 'patient shows concern'. However, one of the 'rules of thumb' reported in the Roter manual states that 'if a decision must be made between categorizing an utterance in an instrumental or affect category, the affect category should be used' (Roter, 1991, p. 3).

Secondly, a category to classify patient's crying is lacking in the RIAS. This affective behaviour was classified in the 'expressing concern or worry' category. Since crying can not truly be considered a verbal utterance it was difficult to determine how many tallies to place in this category. The added affective rating on the 'signs of distress' scale compensated for this shortcoming.

Thirdly, some categories were never or rarely used. The categories which were never used were: 'asking for opinion' and 'asking for reassurance' by the oncologist and 'giving directions/orientation' by patients. The rarely used categories were: 'expressing criticism or disapproval' by both oncologists and patients; 'bids for repetition' by oncologists and patients, and 'asking for services' by patients.

STUDY 2: DISCRIMINANT VALIDITY OF THE RIAS

Method

Samples

Oncological sample: Five gynaecologists (one

Table 1. Inter-rater correlation coefficients (*r*) for physician and patient

(a) Regarding the instrumental (task) clusters and categories (*N* = 25)

Instrumental clusters and categories	*r* physician	*r* patient
Directions:	0.90	*
Orientations and/or instructions	0.90	*
Question-asking (closed and open-ended):	0.96	0.95
Medical condition (closed/open)	1.00/0.88	0.91
Therapeutic regimen (closed/open)	0.83/0.68	0.93
Lifestyle (closed/open)	0.94/0.80	0.72
Psychosocial feelings (closed/open)	0.69/0.89	0.60
Other (closed/open)	1.00/*	1.00

Note: all questions are open-ended for patients

Information-giving:	0.98	0.99
Medical condition	0.99	0.98
Therapeutic regimen	0.98	0.79
Lifestyle, psychosocial feelings	0.91	0.93/0.90
Other	0.94	0.89
Counselling (persuasion):	0.98	**
Medical condition, therapeutic regimen	0.87	**
Lifestyle, psychosocial feelings	1.00	**

(b) Regarding the affective (socio-emotional) clusters and categories (*N* = 25)

Affective clusters and categories	*r* physician	*r* patient
Social behaviour:	0.86	0.97
Personal remarks	0.82	1.00
Friendly jokes (laughter)	0.88	1.00
Approval	0.90	0.69
Compliments	1.00	0.72
Verbal attentiveness:	0.99	0.99
Agreement, showing understanding	0.99	0.99
Paraphrasing, checking	0.93	0.91
Empathy, legitimizing	0.85	*
Showing concern:	0.96	0.94
Concern, worry	0.91	0.95
Reassurance, optimism	0.97	0.79
Negative talk:	0.87	0.76
Disapproval	1.00	0.80
Criticism	*	0.83

* Empty category.
** Provider category only.

Table 2. Inter-rater correlation coeffficients for physician and patient regarding the global affect scales ($N = 25$)

Dimensions	Physician	Patient
Anger/irritation	1.00	0.80
Anxiety/nervousness	1.00	0.84
Dominance/assertiveness	0.88	0.83
Interest/engagement	0.68	0.86
Friendliness/warmth	0.55	0.75
Signs of distress	1.00	0.80

woman, four men), two medical oncologists (one woman, one man), and one (male) urologist recorded initial consultations with cancer patients for the second study. Eligible were patients who had been referred for an initial discussion of possible treatment, and who gave their informed consent. Exclusion criteria were inability to speak, read or write the Dutch language and obvious psychopathology.

Sixty consecutive cancer patients comprised the sample: 12 men and 48 women. Of the 48 women, 39 were gynaecology patients and 9 were medical oncology patients. Of the 12 male patients, 3 were urology patients and 9 were medical oncology patients. The mean age was 54 years (S.D. = 17.9). The gynaecological cancers included cancer of the cervix, vulva, uterus and ovary. The urological cancers were all prostate malignancies. The medical oncology patients were diagnosed with carcinoma of the testis, bladder, liver, pancreas, esophagus, colon, and breast.

During the recorded consultations the following topics were usually discussed: reason for referral to a specialist; the specific aspects of the diagnosis; medical history; findings of the physical examination; the proposed treatment policy; type of medication; side effects; prognosis; consequences of the treatment such as fatigue; consequences for bladder and/or bowel function; vomiting; skin irritation; hair loss; sexuality; fertility; and, work. The physical examination which took place during these consultations was not recorded for technical and privacy reasons.

General practice samples: For the comparison with non-oncological consultations, data from three different general practice samples were available: 329 random-GP consultations; 103 patients diagnosed with hypertension and 65 patients who visited their doctor primarily for a

gynaecological problem (in principle non-malignant).

The mean age of the patients from the random-GP consultations was 49 years (S.D. = 16.6). This sample consisted of 98 men and 168 women (for 86 patients data about age were missing; for 63 patients data about sex were missing). Patients were recruited consecutively. Fifteen general practitioners (two women, 13 men) recorded their consultations.

The mean age of the patients diagnosed with hypertension was 58 years (S.D. = 14.9). This sample consisted of 36 men and 67 women, recruited consecutively. Seventeen male general practitioners were included.

The mean age of the gynaecology population was 36 years (S.D. = 13.4). These patients were recruited consecutively. Seventeen general practitioners (11 women, six men) taped their consultations.

Instrument. The RIAS was used for analyzing consultations. Several categories which were not included in either the instrumental or affective clusters were all comprised in the 'miscellaneous' cluster: asking for reassurance; transitions; bid for repetition; asking for understanding; asking for opinion; asking for services by the patient. Unintelligible utterances were regarded as missing values.

For the oncological sample, a sixth scale, 'signs of distress', was added to the global affective scales for both patients and physicians.

Procedure and analyses. The oncological consultations were taped as part of a larger randomized experimental study investigating the efficacy of providing patients with audiotapes of their oncological consultations. The general practice consultations were collected earlier by researchers from the NIVEL. Their data were made available for comparison with our oncological sample.

Both the oncological consultations and the random GP consultations were analyzed by two coders. The hypertension sample was analyzed by one coder; the gynaecology sample by three coders. The oncological consultations were coded from audiotape. The coding of the general practice consultations took place from videotape, which made it possible to register visual cues as well.

For the overall statistics, the relative contribution of both physicians and patients to the consultation was calculated for all samples. For every consultation, the contribution of physicians and patients was calculated in terms of percentages instead of counts. These percentages were then averaged over all consultations to determine the 'mean percentages per RIAS clusters'. Finally, the 'total instrumental', 'total affective', 'miscellaneous' and 'missing' percentages for both physicians and patients in each sample were added. Also, two-sample t-tests (Levene's Test for Equality of Variances) were performed to determine possible differences between samples, regarding the mean number of utterances per consultation.

In order to determine the discriminant validity of the RIAS, the mean percentages and S.D. of the clusters were calculated. Since we formulated plausible expectations rather than strong hypotheses, two-tailed t-tests were carried out to identify differences in communicative behaviours between the four samples.

Finally, the mean ratings of the global affective scales were computed. Because of the skewness of the data, Mann-Whitney U tests were carried out to identify whether differences in affect between samples were according to the expectations formulated.

To correct for multiple testing, the Bonferroni correction was used. Therefore, the p-value was set on $p < 0.001$. All statistical analyses were performed by using SPSS for Windows 6.0.1.

Results

Overall statistics. From Table 3 it can be seen that the oncological consultation was physician-dominated: oncologists contributed exactly 60% to the consultation whereas patients' contribution was 40%. In the three general practice samples the contribution of physicians and patients was more balanced: general practitioners contributed between 49.2 and 54.3%, and patients between 45.7 and 50.8%.

The mean number of utterances per consultation was 322.8 (S.D. = 137.4) for the oncological sample, 246 (S.D. = 134.6) for the random-GP sample, 284.6 (S.D. = 177.4) for the hypertension sample, and 331.9 (S.D. = 149.2) for the gynaecological sample. The mean number of utterances per consultation was higher for the oncological sample than the random GP sample ($t = 4.05$,

df = 387, $p < 0.001$). There were no differences found between the oncological sample and hypertension sample, and the gynaecological sample.

Differences in communicative behaviours between oncologists and general practitioners. The expectations regarding oncologists' instrumental behaviours were confirmed. Oncologists gave more information than the GPs in the random-GP sample ($t = 7.52$, df = 65.43, $p < 0.001$), the hypertension sample ($t = 9.53$, df = 71.84, $p < 0.001$) and the gynaecology sample ($t = -0.67$, df = 86.39, $p < 0.001$). As expected, oncologists gave less counselling statements than the GPs in the random-GP sample ($t = -13.41$, df = 283.43, $p < 0.001$), the hypertension sample ($t = -12.64$, df = 133.75, $p < 0.001$), and the gynaecology sample ($t = 8.87$, df = 82.88, $p < 0.001$).

With respect to physicians' affective behaviours, both expectations were partially confirmed. Oncologists were more verbally attentive than GPs in the random-GP sample ($t = 4.25$, df = 387, $p < 0.001$), and the gynaecology sample ($t = -5.35$, df = 93.33, $p < 0.001$). No differences between oncologists and GPs in the hypertension sample were found. Further, oncologists expressed more concern than the GPs in the gynaecology sample ($t = -4.28$, df = 97.83, $p < 0.001$). There were no differences found between oncologists and GPs in both the random and hypertension sample.

From Table 3 it can be seen that, apart from the expectations formulated, physicians differed from each other regarding other clusters as well. *Post hoc* comparisons showed that oncologists gave least directive statements as compared to GPs in the random-GP sample ($t = -12.24$, df = 252.95, $p < 0.001$), the hypertension sample ($t = -6.13$, df = 155.15, $p < 0.001$), and the gynaecology sample ($t = 5.68$, df = 76.63, $p < 0.001$).

Oncologists also showed less social behaviour compared to GPs in the random-GP sample ($t = -11.41$, df = 265.91, $p < 0.001$), the hypertension sample ($t = -7.71$, df = 133.16, $p < 0.001$), and the gynaecology sample ($t = 4.86$, df = 79.37, $p < 0.001$). In addition, oncologists expressed less negative talk compared to GPs in the hypertension sample ($t = -4.31$, df = 116.24, $p < 0.001$).

Differences in communicative behaviours between cancer patients and general practice patients: The expectation regarding the instrumental behaviours

Table 3. Mean (%) and S.D. of instrumental and affective clusters physicians and patients for all samples

	Oncological sample (N = 60)		Random-GP sample (N = 329)			Hypertension sample (N = 103)			Gynaecological sample (N = 65)		
	Mean (%)	(S.D.)	Mean (%)	(S.D.)		Mean (%)	(S.D.)		Mean (%)	(S.D.)	
Instrumental clusters physician:											
Directions	2.6	(1.4)	6.2	(4.0)	*	4.8	(3.1)	*	6.2	(4.8)	*
Questions	7.8	(5.5)	7.1	(3.9)		5.4	(2.9)		7.1	(3.0)	
Information	21.6	(10.9)	10.8	(5.9)	*	7.5	(4.7)	*	14.3	(5.6)	*
Counselling	1.2	(1.3)	4.7	(3.8)	*	6.5	(3 9)	*	5.1	(3.4)	*
Total instrumental physician	33.2		28.8			24.2			32.7		
Instrumental clusters patient:											
Directions	0.0	(0.1)	1.0	(1.4)	*	0.5	(0.7)	*	0.4	(0.5)	*
Questions	1.8	(1.5)	1.7	(1.5)		2.6	(1.8)		2.0	(1.3)	
Information	15.2	(8.9)	23.6	(8.0)	*	25.1	(8.3)	*	23.4	(7 0)	*
Counselling	Provider category only		Provider category only			Provider category only			Provider category only		
Total instrumental patient	17.0		26.3			28.2			25.8		
Affective clusters physician:											
Social behaviour	1.1	(1.0)	3.6	(3.0)	*	3.8	(3.3)	*	3.1	(3.3)	*
Verbal attentiveness	16.7	(6.1)	13.2	(5.9)	*	14.4	(6.0)		11.8	(4.1)	*
Showing concern	3.1	(2.3)	2.0	(2.0)		2.2	(1.8)		1.6	(2.2)	*
Negative talk	0.0	(0.1)	0.1	(0 4)		0.3	(0 6)	*	0.1	(0.2)	
Total affective physician	20.9		18.9			20.7			16.6		
Miscellaneous	5.6	(4.3)	4.9	(2.7)		3.7	(2.6)		4.2	(2.2)	
Missing	0.3	(0.6)	0.8	(2.0)		0.6	(2.3)		0.8	(1.2)	
Affective clusters patient:											
Social behaviour	1.3	(1.4)	4.1	(3.9)	*	4.3	(5.2)	*	3.2	(3.2)	*
Verbal attentiveness	15.6	(8.0)	9.1	(4.1)	*	9.6	(3.9)	*	11.4	(4.0)	*
Showing concern	2.4	(2.4)	3.9	(3.3)		4.2	(2.3)	*	2.4	(2.2)	
Negative talk	0.1	(0.2)	0.2	(0.7)		0.4	(0.8)	*	0.1	(0.2)	
Total affective patient	19.4		17.3			18.5			17.1		
Miscellaneous	1.9	(1.4)	2.3	(1.6)		3.3	(2.3)		1.6	(1.3)	
Missing	1.7	(1.7)	1.0	(1.8)		0.8	(1.5)		1.2	(2.0)	
Total:	100		100			100			100		

All GP-samples were tested against the oncological sample; * $p < 0.001$.

of cancer patients was not confirmed. Cancer patients did not ask more questions than patients in the random-GP sample ($t = 0.40$, df $= 387$, $p = 0.691$), hypertensive patients ($t = -2.89$, df $= 161$, $p = 0.004$), and gynaecology patients ($t = 0.97$, df $= 123$, $p = 0.332$).

The expectation regarding the affective behaviours of cancer patients was also not confirmed. Rather, patients in the hypertension sample showed more concern than cancer patients ($t = -4.56$, df $= 161$, $p < 0.001$). No differences between cancer patients and patients in the random-GP sample and gynaecology sample were found.

From Table 3, other differences between the patient samples can be noticed. *Post hoc* comparisons indicated that cancer patients were least directive as compared to patients in the random-GP sample ($t = -11.51$, df $= 349.47$, $p < 0.001$), hypertensive patients ($t = -5.91$, df $= 110.72$, $p < 0.001$), and gynaecology patients ($t = 6.45$, df $= 72.08$, $p < 0.001$). Cancer patients also gave less information than patients in the random-GP sample ($t = 7.23$, df $= 387$, $p < 0.001$), hypertensive patients ($t = -7.00$, df $= 161$, $p < 0.001$), and gynaecology patients ($t = 5.68$, df $= 112.61$, $p < 0.001$).

Like oncologists, cancer patients showed less social behaviour than patients in the random-GP sample ($t = -9.99$, df $= 246.80$, $p < 0.001$), hypertensive patients ($t = -5.62$, df $= 126.78$, $p < 0.001$), and gynaecology patients ($t = 4.33$, df $= 89.29$, $p < 0.001$). However, cancer patients were more verbally attentive than the patients in the random-GP sample ($t = 6.22$, df $= 64.74$, $p < 0.001$), hypertensive patients ($t = 5.48$, df $= 75.37$, $p < 0.001$), and gynaecology patients ($t = -3.68$, df $= 86.18$. $p < 0.001$). They also expressed less negative talk as compared to hypertensive patients ($t = -3.43$, df $= 128.19$, $p < 0.001$). There were no differences found regarding expression of negative talk between cancer patients, random GP-patients and gynaecology patients.

Differences in global affect between oncologists and general practitioners: Table 4a shows the median and range of the global affective scales of physicians for all samples. The expectations regarding oncologists' global affect were not confirmed. Rather, physicians in both the random-GP sample and the gynaecology sample had a higher rating on the interest/engagement scale compared to the oncologists (random-GP sample: median 5 *vs.* 4, $Z = -5.9$, $p < 0.001$; gynaecology sample:

median 5 *vs.* 4, $Z = -6.2$, $p < 0.001$). There was no difference between oncologists and the GPs in the hypertension sample regarding this affective scale. Also contrary to our expectation, physicians in the random-GP sample and gynaecology sample had a higher rating on the friendliness/warmth scale as compared to the oncologists (random-GP sample: median 5 *vs.* 4, $Z = -5.0$, $p < 0.001$; gynaecology sample: median 5 *vs.* 4, $Z = -6.3$, $p < 0.001$). There was no difference regarding oncologists and GPs in the hypertension sample.

Post hoc comparisons indicated that, besides the expectations formulated, all other differences in affect between physician samples were statistically significant.

Differences in global affect between cancer patients and general practice patients: Table 4b shows the median and range of the global affective scales of patients for all samples. The expectation concerning cancer patients' ratings for the anxiety/nervousness scale was not confirmed. Hypertensive patients had higher ratings compared to cancer patients (median 4 *vs.* 1.5, $Z = -8.3$, $p < 0.001$). There was no difference regarding this type of affect for cancer patients and both patients in the random GP-sample and gynaecology patients.

Apart from the expectations tested, it can be seen from Table 4b that cancer patients differed from general practice patients regarding other affective dimensions. *Post hoc* comparisons showed that, besides the expectation formulated, most other differences in affect between patient samples were statistically significant. There were no differences regarding the anger/irritation scale between cancer patients and both patients in the random GP-sample and gynaecology patients. There was also no difference regarding the dominance/assertiveness scale between cancer patients and hypertensive patients.

DISCUSSION

In this paper, we presented two studies in which the psychometric properties of the Roter Interaction Analysis System (RIAS) were investigated in an oncological setting. We undertook the first study to examine the feasibility, inter rater-reliability and content validity of the RIAS. The second study was aimed at evaluating the discriminant validity of the RIAS.

Table 4. Median and range of the global affective scales of physicians for all samples

	Oncological sample (N = 60)		Random-GP sample (N = 329)			Hypertension sample (N = 103)			Gynaecological sample (N = 65)		
	Median	Range	Median	Range		Median	Range		Median	Range	
(a) Affective scales physicians											
Anger/irritation	1.0	[1–1]	1.0	[1–5]		5.0	[1–6]	*	1.0	[1–3]	
Anxiety/nervousness	1.0	[1–1]	1.0	[1–5]		5.0	[2–6]	*	1.0	[1–3]	
Dominance/assertiveness	4.0	[3–5]	5.0	[2–6]	*	2.0	[1–5]	*	5.0	[3–6]	*
Interest/engagement	4.0	[2–6]	5.0	[1–6]	*	4.0	[1–6]		5.0	[4–6]	*
Friendliness/warmth	4.0	[2–5]	5.0	[1–6]	*	4.0	[1–6]		5.0	[3–6]	*
Signs of distress	1.0	[1–2]	Oncological rating only			Oncological rating only			Oncological rating only		
(b) Affective scales patients											
Anger/irritation	1.0	[1–2]	1.0	[1–5]		5.0	[1–6]	*	1.0	[1–4]	
Anxiety/nervousness	1.5	[1–5]	2.0	[1–6]		4.0	[1–6]	*	1.0	[1–5]	
Dominance/assertiveness	3.0	[1–5]	4.0	[2–6]	*	3.0	[1–6]		4.0	[2–6]	*
Interest/engagement	4.0	[2–5]	5.0	[2–6]	*	5.0	[2–6]	*	5.0	[4–6]	*
Friendliness/warmth	3.0	[2–5]	4.0	[2–6]	*	4.0	[1–6]	*	5.0	[3–6]	*
Signs of distress	2.0	[1–5]	Oncological rating only			Oncological rating only			Oncological rating only		

With respect to the scales measuring anger/irritation, anxiety/nervousness, and signs of distress, 'normal or average' affect is rated as 1 or 2. Regarding dominance/assertiveness, interest/engagement, friendliness/warmth, a rating of 3 or 4 is considered 'normal or average' affect.
All GP-samples were tested against the oncological sample; $* p < 0.001$.

The results of the first study showed that the coding of taped oncological consultations was more time consuming than indicated by Roter (1991). The Roter manual reports that coding time is about double the duration of the consultation. In our study, coding time was about five times the duration of the consultation. This finding is consistent with a Dutch study by Bensing (1991a,b), for which the average coding time was even higher. It could be hypothesized that translation of Dutch utterances into the English RIAS categories took extra time. Another explanation for the longer coding time in our study could be the oncological setting. In almost all cases there were more than two people present, which made communication more difficult to distinguish. The fact that Bensing, using GP-samples, also found coding more time consuming challenges this explanation. Getting more familiarized with the RIAS manual over time, however, seemed to reduce coding time. Nevertheless, in general, the process of keeping tallies for each utterance is time consuming and attention demanding. In all likelihood, a computerized version of the RIAS would enhance the feasibility of coding (see Appendix A).

The inter-rater correlations of the RIAS proved to be high. This holds true for the individual RIAS categories, the RIAS clusters, and to a lesser extent, the global affective scales. Since the affective scales are subjective in nature and thus can not be classified according to fixed coding rules, a lower reliability can be expected. All previous studies using the RIAS reported a high inter-rater reliability. Because of the skewness of the data concerning the under-utilized categories, most researchers reported Pearson correlations for frequently used categories only. For example, Hall et al. (1987) reported inter-rater reliability for 17 out of 62 categories, Roter et al. (1991) for 21 physician categories and 11 patient categories, and Ford et al. (1995) for 15 categories. In our study we calculated the reliability for all categories by means of the Spearman Rank Correlation Coefficient, which is less sensitive for deviations from normality. For reasons of comparability with other studies, however, we also calculated Pearson Correlations for these categories. It appeared that both correlation coefficients showed high reliability although the Spearman Rank Correlations were slightly lower on average. We decided that, from a methodological point of view, Spearman Correlations were more appropriate to report.

It can be questioned however, whether calculating correlation coefficients for categories and clusters, is the way to determine the inter-rater reliability of the RIAS. It would probably be more suitable to determine the congruence in categorizing per utterance and afterwards calculating a coefficient such as Cohen's Kappa. This is only possible, however, if information about the sequence of the categorizing process is known. At the moment, a computerized version of the RIAS is available in which sequential information is kept (Appendix A). Still, one problem remains. This concerns the fact that there are two decisional tasks when applying the RIAS. One task means deciding what fragment of the consultation comprises exactly one utterance. The other task relates to deciding in which category an utterance should be classified. Our study results showed that the congruence between the number of utterances distinguished is high. However, the number of utterances within a consultation distinguished by the two raters is not exactly equal. This is, however, necessary for determining the inter-rater reliability following the method described above. Therefore, predefining the content of each utterance (and therefore the number) is a prerequisite for determining Cohen's Kappa. The agreement about which utterances (number and content) are to be distinguished can then be determined separately.

Another topic that relates to the inter-rater reliability is the analysis on cluster level. Although Roter distinguishes 62 categories in her original system (1991), these categories are not the relevant units of analyses in studies applying the RIAS (Hall et al., 1987, 1994a,b; Roter et al., 1991; Ford et al., 1995). The categories are always merged into clusters. These clusters are then used to describe the content of the conversation and the relations with other (outcome) variables. Therefore, it seems more meaningful to determine the inter-rater reliability per cluster than per category. This will result in even higher correlation coefficients and also solve the problem of handling rarely used categories. Nevertheless, a complicating factor is that different researchers distinguish different RIAS clusters. For reasons of comparability it would be convenient that standard clusters are used. On the other hand, a standardized instrument could be less relevant for certain research aims. Standardization of clusters could undermine the flexible nature of the RIAS, and thus limit the exploration of new or alternative theoretical

approaches. These aspects have to be weighed carefully.

Because of our dissatisfaction with the correlations as an indicator of the inter-rater reliability, we also calculated difference scores per category for the two raters. These findings indicated that there were considerable differences between the raters. Especially, for the categories 'physician transitions' and 'physicians giving information about lifestyle and psycho-social feelings', there were instances of large differences between the coders (difference scores of 22 and 18). Both the Pearson and Spearman correlations varied between 0.83 and 1.00 for categories which showed the highest difference scores. This finding strengthens our opinion that these correlations can be misleading when calculating the reliability of the RIAS.

Regarding the content validity of the RIAS, all utterances of both oncologists and cancer patients could be classified. Nevertheless, coding conflicts occurred concerning classification of affective utterances expressing concern and optimism categories. However, since these two categories comprise one affective cluster, 'expressing concern', it is not problematic when analyzing on cluster level. The inclusion of a separate category for patient's crying might be useful, although the added global affective scale 'signs of distress' could also reflect this emotional type of behaviour. It was remarkable, however, that scores on the 'signs of distress' scale were rather low. This challenges the need for an extra communication category. Also, an additional category for 'crying' would not solve the problem of determining how many tallies to place in this category.

When looking at the global affective scales it is striking that there is very little variance. Moreover, there is no variance concerning the scales 'anger/irritation' and 'anxiety/nervousness' of the physician: all ratings are exactly the same. This seriously challenges the usefulness of these scales within oncology samples. However, in a previous study by Bensing (1991a,b), the global affective ratings proved to have the greatest predictive power in quality assessments of care in a general practitioners sample. It is plausible that these results can partly be attributed to the use of video, instead of audiotaped consultations in the GP-samples; previous studies have reported that 55% of affective communication is transferred by visual cues, eye-contact, body positioning, etc. (Bensing, 1991a,b). Besides, the affective scales

appear to have discriminant validity in this study with respect to GP-samples. Therefore, it seems that one has to be careful with omitting these scales altogether.

An oncology-specific modification of the RIAS could consist of a three-way coding system, with additional communication categories for the patient's partner and/or other relatives. The RIAS manual states that 'if a third party is present, the rater should briefly describe the role that this third person plays, including length of time present and degree of involvement' (Roter, 1991). However, in our oncological sample almost all patients were accompanied by a family member. This third person was usually present during the entire consultation (except for the physical examination), and often contributed to the consultation to a considerable degree. A revised three-way RIAS system has already been developed for a paediatric setting (Wissow *et al.*, 1994). Likewise, Butow *et al.* (1995) have developed a computerized interaction analysis system which allows for coding of a third party present.

Some RIAS categories were never or rarely used. It could be hypothesized that the 'asking for opinion by the physician' category is more appropriate for a general practice setting. The same probably holds true for 'giving directions' and 'asks for services' by the patient. These categories seem less relevant for an oncological setting. However, for comparison purposes these categories should not be excluded. Also, these categories themselves carry significance. They could, for instance, function as important markers for change towards a more 'patient-centered' model of medical care.

Caution is needed when interpreting the results of the second study. First of all, the four samples are quite different in terms of age and sex. The communication differences found between the oncological sample and the three general practice samples could be attributed to these socio-demographic variations. Second, our oncological patients can not be considered a representative sample in terms of sex, age, and tumour sites. Most cancer patients (39/60) were recruited from the gynaecology outpatient clinic; women were therefore over-represented in the oncological sample.

The results of the second study showed that most expectations regarding differences in physician communication were confirmed. Oncologists gave more information and were less persuasive

than GPs in all samples. Oncologists were also more verbally attentive in two out of three comparisons, and expressed more concern in one out of three comparisons. *Post hoc* comparisons yielded additional differences between physicians. As mentioned earlier, most cancer patients were accompanied by a family member who contributed to the consultation to a considerable degree. Information-giving by the oncologists could be a result of questions asked by a relative. Since we chose to only analyze doctor–patient communication, this assumption can not be verified. The fact that the physical examination was not recorded in our oncology sample could have resulted in an under-representation of directive statements by oncologists. Oncologists also expressed less social behaviours as well as negative talk. In all likelihood, the often longer existing and on-going relationship between GPs and patients has led to a more egalitarian relationship. This type of relation probably facilitates expression of these behaviours, both positive and negative affect.

The expectations regarding patient communication were not confirmed. Cancer patients did not ask more questions and did not express more concern as compared to general practice patients. The fact that all cancer patients were referred by another physician could perhaps explain why they did not ask more questions; these patients already received some information about their diagnosis and treatment from their referring physician. Also, oncologists were more informative than general practitioners which may have resulted in less question-asking by cancer patients. Patients' experiences with their illness may also have an effect on question-asking. As compared to newly diagnosed cancer patients, patients suffering from a chronic disease may be more assertive in question-asking since they have the benefit of a long-gained experience with their condition. Their achieved knowledge may function as a 'cognitive frame' from which they derive additional questions. This may account for the higher number of questions from hypertensive patients. Finally, since the physical examination was not recorded, cancer patients' question-asking was possibly under-represented.

As mentioned earlier, cancer patients did not show more concern as compared to general practice patients. *Post hoc* comparisons did show that, like oncologists, cancer patients were more verbally attentive. This behaviour could be in response to the highly informative character of the oncological consultation. The fact that general practice patients expressed more social behaviours and negative talk could result from the more egalitarian relationship that these patients have with their physician. As mentioned earlier, consultations with a GP often occur within a longer-existing relation as compared to a first consultation with an oncologist.

In general, the affect ratings of the general practice samples were only slightly higher in comparison with the oncological sample. The hypertension sample, however, had noticeable higher ratings regarding the anger/irritation and anxiety/nervousness scales, as compared to the other three samples. In the oncological sample, all affective scales showed 'normal or average' affect. This finding may be counter-intuitive, although it should be remembered that these oncological consultations concerned treatment options and can be considered less emotional than for instance 'bad news' consultations.

Another explanation for the average scores on these affective scales in the oncology sample could be the nature of these scales. If a rater has to judge affective behaviour on a Likert-type scale, other consultations will function as a frame of reference. In most instances, raters only code consultations of one particular sample (e.g. oncological, general practice). So, it can be expected that the mean scores are about average. This explanation challenges the use of these scales for detecting differences between samples. Finally, video instead of audiotaping oncological consultations would perhaps yield different global affect ratings. As mentioned earlier, video recordings enable registration of both vocal and visual cues.

The results of the first study show that the feasibility and content validity of the RIAS proves to be satisfactory in an oncological setting. Also, the inter-rater reliability appears to be high, although it is highly probable that a more advanced method for determining the inter-rater reliability would yield less positive results.

The results of the second study indicate that the RIAS is able to discriminate between communicative behaviours in an oncological sample and three general practice samples. Future studies, using samples which are matched in terms of, for example age and sex, would have to demonstrate whether the communication differences reported above can be replicated.

ACKNOWLEDGEMENTS

APPENDIX A: COMPUTERIZED VERSION OF THE RIAS

The RIAS appears to be a highly reliable tool in studying doctor–patient communication. However, it has obvious disadvantages. One of them concerns (a) the imperfect ergonomic properties of the coding task. The task is highly attention-demanding and therefore error-prone.

While listening to the audiotape, the rater has to assign each utterance (communication unit) to one of the 62 communication categories of the RIAS. Classifications are noted down on paper by keeping tallies for each category. This apparently simple task, however, appears to be very attention demanding. It is easy to lose hold of the right utterance to be coded, when interrupted by, for example, a telephone call, a knock on the door or even an intervening thought. On the basis of the tallies it is practically impossible to catch up. One has to start anew. This is a time-consuming process that demands a lot of discipline. Another problem with this method is the extra step that is required by entering the paper and pencil data into the computer. This takes both time and introduces a possible source of errors.

A second disadvantage refers to (b) the loss of the sequential information concerning the interaction. Therefore, a computerized version of the RIAS has been developed which both is more user-friendly and stores sequential information. The way of coding as propagated by Roter yields several meaningful categories of utterances that have proven to relate to important outcome variables as for instance patient satisfaction (Ong *et al.*, 1995). However, one important aspect of information is lost: the sequence in which utterances follow one another. Probably relevant aspects of the doctor–patient interaction such as the way in which doctors react to patients' expressions of concern, or, their need for information can only be studied if sequential information is provided.

The ergonomic properties of the coding task together with the need for sequential information has led us to develop an adapted, computerized version of the RIAS coding system. With this program, classification of utterances can be done directly on computer. As a result, the extra step of entering paper and pencil data into the computer is omitted. Also, sequential information is kept. Moreover, because the ten last classifications are constantly visible on the screen there is direct feedback about the ongoing conversation. This provides an important memory aid with respect to which utterance has to be coded next. As a consequence, the task becomes less attention demanding and therefore less error-prone. By giving the opportunity to save the content of the last coded utterance, an additional memory aid is provided for shorter and longer breaks. Next to their function as a memory aid, the 10 last coded utterances also provide a kind of feedback, which makes the task more reinforcing.

Because the program is written in Turbo Pascal for Windows, a Window environment is needed to run the program. Data are saved in ASCII-format. This format is suitable to be used in statistical packages like SPSS. Moreover, an assistant program is developed to transform the data into the most manageable format for SPSS-analyses.

Several other facilities are provided such as the administration of personal data (patient's date of birth, gender, date of consultation, doctor's, patient's and coder's identification), an in-between back-up option and the option to give global ratings for doctor and patient on five affective dimensions, which are also part of the original RIAS.

To conclude, the computerized version of the RIAS[1] coding system is an improvement in comparison to the old system for several reasons, namely:

1. By providing a constant memory aid the classification task becomes less attention demanding.
2. Sequential information is provided.
3. Categorizing data are entered directly into the computer.
4. Direct feedback makes coding more reinforcing.
5. It provides the opportunity to save the content of the last coded utterance.
6. Data can be used directly in statistical packages like SPSS.
7. It provides facilities to store background information.

NOTES

1. The adapted version of the RIAS is available against cost-price. For information contact the second author (E-mail: m.r.visser@amc.uva.nl).

REFERENCES

Bensing, J.M. (1991a) Doctor–patient communication and the quality of care. *Soc. Sci. Med.* **32,** 1301–1310.

Bensing, J.M. (1991b) Doctor–patient communication and the quality of care. An observation study into affective and instrumental behavior in general practice. Dissertation. NIVEL, Utrecht.

Blanchard, C.G., LaBreque, M.S., Ruckdeschel, J.C. and Blanchard, E.B. (1988) Information and decision-making preferences of hospitalized adult cancer patients. *Soc. Sci. Med.* **27,** 1139–1145.

Butow, P.N., Dunn, S.M., Tattersall, M.H.N. and Jones, Q.J. (1995) Computer-based interaction analysis of the cancer consultation. *Br. J. Cancer* **71,** 1115–1121.

Carter, W.B., Inui, T.S., Kukull, W. and Haigh, V. (1982) Outcome-based doctor–patient interaction analysis: II. Identifying effective provider and patient behavior. *Med. Care* **20,** 550–566.

Castejón, J., López-Roig, S., Pastor, M.A., Picó, C., Reig, M.T., Rodríguez-Marín, J. and Terol, C. (1993) *Cancer patients: Health information and quality of life*. The 7th Conference of the European Health Psychology Society. Abstract, p. 18.

Chaitchik, S., Kreitler, S., Shaked, S., Schwartz, I. and Rosin, R. (1992) Doctor–patient communication in a cancer ward. *J. Cancer Educ.* **7,** 41–54.

Eussen, G.P.H.E., Borgers, M.Th.A. and Visser, A.Ph. (1992) Een observatiesysteem voor de analyse van de inhoud van de communicatie tussen kankerpatienten en specialisten. *Gedrag Gezondh.* **20,** 102–106.

Fallowfield, L.J., Baum, M. and Maguire, G.P. (1986) Effects of breast conservation on psychological morbidity associated with diagnosis and treatment of early breast cancer. *Br. Med. J.* **293,** 1331–1334.

Fallowfield, L.J., Hall, A., Maguire, G.P. and Baum, M. (1990) Psychological outcomes of different treatment policies in women with early breast cancer outside a clinical trial. *Br. Med. J.* **301,** 575–580.

Fallowfield, L., Ford, S. and Lewis, S. (1995) No news is not good news: information preferences of patients with cancer. *Psycho-Oncology* **4,** 197–202.

Ford, S., Fallowfield, L. and Lewis, S. (1995) Doctor–patient interaction in oncology. *Soc. Sci. Med.* **42,** 1511–1519.

Hall, J.A., Roter, D.L. and Katz, N.R. (1987) Task versus socio-emotional behaviors in physicians. *Med. Care* **25,** 399–412.

Hall, J.A., Irish, J.T., Roter, D.L., Ehrlich, C.M. and Miller, L.H. (1994a) Satisfaction, gender, and communication in medical visits. *Med. Care* **32,** 1216–1231.

Hall, J.A., Irish, J.T., Roter, D.L., Ehrlich, C.M. and Miller, L.H. (1994) Gender in medical encounters: an analysis of physician and patient communication in a primary care setting. *Health Psychol.* **13,** 384–392.

Inui, T.S., Carter, W.B., Kukull, W.A. and Haigh, V.H. (1982) Outcome-based doctor–patient interaction analysis. 1. Comparison of techniques. *Med. Care* **20,** 535–549.

Kaplan, S.H., Greenfield, S. and Ware, J.E. (1989) Assessing the effects of physician–patient interactions on the outcomes of chronic disease. *Med. Care* **27,** S110–S127.

Molleman, E., Krabbendam, P.J., Annyas, A.A. *et al.* (1984) The significance of the doctor–patient relationship in coping with cancer. *Soc. Sci. Med.* **18,** 475–480.

Ong, L.M.L., Haes, J.C.J.M., de Hoos, A.M. and Lammes, F.B. (1995) Doctor–patient communication: a review of the literature. *Soc. Sci. Med.* **40,** 903–918.

Roter, D.L., Hall, J.A. and Katz, N.R. (1987) Relations between physicians' behaviors and analogue patients' satisfaction, recall, and impressions. *Med. Care* **25,** 437–451.

Roter, D.L., Hall, J.A. and Katz, N.R. (1988) Patient–physician communication: a descriptive summary of the literature. *Patient Educ. Couns.* **12,** 99–119.

Roter, D.L. (1991) *The Roter Method of Interaction Process Analysis* (RIAS Manual).

Roter, D.L., Lipkin, M. and Korsgaard, A. (1991) Sex differences in patients' and physicians' communication during primary care medical visits. *Med. Care* **29,** 1083–1093.

Siminoff, L.A. and Fetting, J.H. (1991) Factors affecting treatment decisions for a life-threatening illness: the case of medical treatment of breast cancer. *Soc. Sci. Med.* **32,** 813–818.

Sutherland, H.J., Llewellyn-Thomas, H.A., Lockwood, G.A. *et al.* (1989) Cancer patients: their desire for information and participation in treatment decisions. *J. R. Soc. Med.* **82,** 260–263.

Waitzkin, H. (1985) Information giving in medical care. *J. Health Soc. Behav.* **26,** 81–101.

Wasserman, R.C. and Inui, T.S. (1983) Systematic analysis of clinician–patient interactions: A critique of recent approaches with suggestions for future research. *Med. Care* **XXI,** 279–293.

Wissow, L.S., Roter, D.L. and Wilson, M.E.H. (1994) Pediatrician interview style and mothers' disclosure of psychosocial issues. *Pediatr.* **93,** 289–295.