Psychological Medicine (2010), **40**, 467–476. © Cambridge University Press 2009
doi:10.1017/S0033291709990730

ORIGINAL ARTICLE

The evaluation of training in patient-centred interviewing skills for psychiatric residents

M. Rimondini*, L. Del Piccolo, C. Goss, M. Mazzi, M. Paccaloni and C. Zimmermann

Department of Medicine and Public Health, Section of Psychiatry and Clinical Psychology, University of Verona, Verona, Italy

Background. Communication skills are considered 'core skills' in the curriculum of psychiatry but studies evaluating the effectiveness of a time-limited training course in interviewing skills in psychiatry have remained rare. The aim was to assess the effectiveness of training in patient-centred interviewing on the interview performance of psychiatric residents.

Method. Psychiatric residents ($n=10$) each interviewed 12 different anonymized standardized patients (SPs), eight before and another four after training. SPs simulated psychiatric out-patients who attended for a first visit to the psychiatric out-patient clinic. The consultations were videotaped, transcribed and coded with a classification scheme developed for psychiatric consultations from which an interview performance index was derived. An interrupted time-series design and a segmented regression analysis with multilevel analysis explored the performance trend within the series of consultations.

Results. The regression model evidenced a horizontal slope at pre- and post-training, with a significant level change. These findings excluded the presence of a practice effect and indicated a significant effect of training. Performance variability between and within residents over the series of consultations increased at post-training.

Conclusions. The training improved patient-centred interviewing performance. More post-training exercise time and supervised practice are necessary to establish consistent performance patterns at a higher skill level.

Received 10 March 2008; Revised 5 June 2009; Accepted 15 June 2009; First published online 23 July 2009

Key words: Interviewing skills, learned-centred training, patient-centred communication, psychiatry.

Introduction

The biopsychosocial model of illness has widened the psychiatrist's set of clinical goals. Psychiatrists are asked to reconcile biopsychosocial data gathering with the development of a trusting relationship and a shared decision-making process and to learn the patient-centred interviewing skills that facilitate achieving these goals. Open-ended inquiry, active listening in terms of reflections, checking, summarizing, clarifying and patient-orienting expressions, together with relationship-building skills such as emotion-seeking questions and emotion-focusing statements improve the information-gathering process and contribute to patient satisfaction and participation and thus to better treatment adherence and health status in different medical settings (Stewart *et al*. 2000; Bodenheimer *et al*. 2002).

Biopsychosocial information gathering and building rapport are particularly important for first consultations in psychiatry for which high treatment drop-out rates (Percudani *et al*. 2002; Rossi *et al*. 2002; Young *et al*. 2002), poor treatment adherence (Nosé *et al*. 2003), low patient satisfaction (Morgan, 1999) and lack of patient involvement (Goss *et al*. 2008) are reported. Patient-centred communication skills would be helpful resources to deal with such qualitative shortcomings.

The importance of enhancing psychiatrists' communication skills is widely recognized. Healthcare organizations such as the Academy of Royal Colleges, the World Psychiatric Association and the American Board of Psychiatry and Neurology have introduced communication skills as 'core skills' in the curriculum of psychiatry (Walton & Gelder, 1999; Swick *et al*. 2006). Programmes on doctor–patient communication are being extended to the psychiatric residency years, and a wide range of assessment tools has been developed (Swick *et al*. 2006).

However, studies that assess the efficacy of time-limited interviewing skill training courses have remained rare in psychiatry, compared with other

* Address for correspondence: M. Rimondini, Ph.D., Dipartimento di Medicina e Sanità Pubblica, Università di Verona, Policlinico G.B. Rossi, Piazzale L.A. Scuro 10, 37134 Verona, Italy.
(Email: michela.rimondini@univr.it)

medical specialities. Three early studies, based on quantitative analyses of word-by-word doctor talk during the consultation, evaluated the effects of a training course on psychiatrists' use of specific patient-centred interviewing techniques (Maguire *et al*. 1984; Lieberman *et al*. 1989; Harrison & Goldberg, 1993). Findings suggested that training in interview skills increased desirable interviewing behaviours, although some skills were more amenable to learning than others. For example, checking and summarizing increased while acknowledging and responding to patients' emotions did not. Undesirable behaviours reducing patients' communicative flow such as asking multiple or closed-ended questions proved resistant to change. Such observations present a challenge to communication research in psychiatry, which only recently has been taken up again (Nuzzarello & Birndorf, 2004; Rimondini *et al*. 2006).

The present study takes this stagnant area of research on the evaluation of interview skills in psychiatry a step further. The study aims to add methodologically sound and updated evidence by applying a design and advanced methods of evaluation fitting the complexity of factors involved. To assess the effectiveness of time-limited communication training for second-year residents in psychiatry in producing measurable changes in interviewing behaviour we used an interrupted time-series (ITS) design (Campbell & Stanley, 1966) with a segmented regression and multilevel analysis to take into account the individual training effect of each participant. ITS is an alternative to randomized clinical trials and is considered the best design to evaluate longitudinal effects of time-limited interventions (Ramsey *et al*. 2003).

The hypothesis was that training would increase patient-centred interviewing skills and/or decrease the use of doctor-centred expressions which reduce patient participation and involvement.

Method

The Verona communication skills training in psychiatry

The course aims to improve residents' interviewing performance during a first psychiatric consultation in out-patient community-based care. Residents learn how to collect reliable information and to establish a relationship with the patient within the time constraints that characterize first psychiatric out-patient visits at most mental health services by using patient-centred interviewing skills.

The course consists of four consecutive weekly small group sessions of 4 h each, with two facilitators (M.R. and C.G.) experienced in learner-centred teaching (Kurtz *et al*. 1998; Joyner & Young, 2006). Teaching methods used are feedback on the videotaped consultations of each participating resident, analyses and exercises regarding the transcripts of these consultations, critical incident reports, and role-play with video feedback. Psychiatrists are taught a reasoned and systematic use of open inquiry, patient-orienting and patient-supporting statements, and active listening skills. They are expected to reduce behaviours that inhibit patients' comunicative flow, such as closed-ended questions, instructions and suggestions and disapproving comments.

Subjects

Psychiatric residents

All 10 psychiatric residents (four females, six males), aged between 28 and 44 years, in their second year of the 4-year postgraduate course in psychiatry of the Medical School of the University of Verona took part. This school accepts residents after a careful selection process among applicants from all over Italy. The selection is based on the academic curriculum and a written admission test on basic knowledge in medicine and psychiatry. In the first 2 years residents' clinical activities in the psychiatric ward, the mental health centre and the out-patient clinic are carried out under close supervision. During the second year they may start to see out-patients but have to report back to their senior colleagues.

Simulated patients

Female ($n = 10$) and male ($n = 10$) actors, aged between 20 and 65 years, from a School of Theatre acted as unannounced psychiatric patients at their first visit at the out-patient clinic of the Community Based Mental Health Service of South Verona. Half of each gender had to present either an anxiety or mood disorder, disorders which characterize the great majority of the psychiatric out-patient population attending the clinic. The actors were paid.

Role scripts

For each of the 20 standardized patients (SPs) a different personal role script was created. The scripts were standardized and contained the same set of information items (although of different content) to collect by the interviewer: five sociodemographic characteristics (name, age, education, current profession, marital status), four symptoms of either anxiety or mood disorder [defined by DSM-IV diagnostic criteria], with information on onset and course of

https://doi.org/10.1017/S0033291709990730 Published online by Cambridge University Press

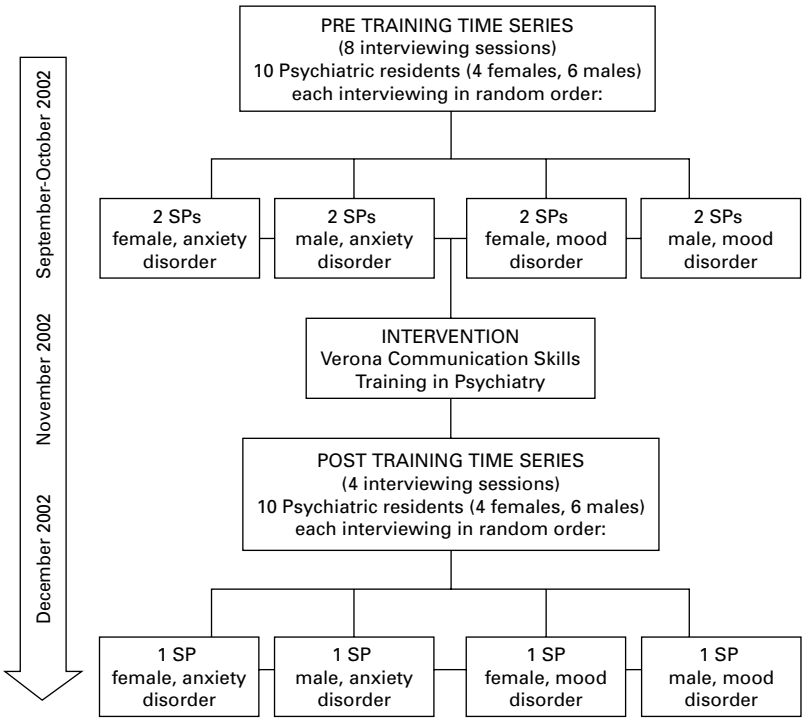

Fig. 1. Flow chart of the study design and data collection procedure. SP, Standardized patient.

symptoms and their impact on functioning and quality of life, the reason for consultation, patient's theory regarding symptoms, treatment expectation, one stressful life event, one severe social problem and information on the patient's social support network. All SPs had been coached individually and had played their role at least twice with one of the authors as interviewer and with feedback by another author after each role-play.

Procedure

The study was approved by the local ethics committee. Residents and SPs gave written, informed consent to be videotaped but were blind to the design of the study. Residents ignored that patients were SPs and SPs that there was a training intervention. The setting of data collection was made as realistic as possible. In September and October 2002, at pre-training (T1), each of the residents interviewed in the out-patient clinic four different female and four different male SPs in random order, half of each gender either presenting an anxiety or mood disorder. Residents had been told by the trainers that these patients had accepted to be videotaped when fixing their out-patient appointment. The training took place in November. In December 2002, at post-training (T2), residents interviewed two new male and female SPs, one of each gender either presenting an anxiety or mood disorder (see Fig. 1). The task of the residents at pre- and post-training was to spend 30 min collecting all the information necessary to understand the SP's current problems and to establish a collaborative relationship. They were not asked to express a diagnosis or suggest a treatment.

Over the study period residents continued to carry out their clinical duties as usual but were encouraged, as part of the training course, to apply the taught skills in their daily clinical practice. Only after the conclusion of the study were the residents informed that the interviewed patients were actors.

The videotaped interviews were fully transcribed. The talk of the residents was divided into units of speech according to criteria proposed by Butow *et al*. (1995). The interviews were equally and randomly attributed to two trained raters who coded the speech units independently using the Verona Psychiatric Interview Classification System (VR-PICS). The raters were blinded to the purposes of the study and to whether the interviews had been recorded during the pre- or post-training sessions.

Measures

VR-PICS

The VR-PICS is a modified version of the Verona Medical Interview Classification System (VR-MICS) (Del Piccolo *et al*. 2004, 2005), adapted to the psychiatric setting (Rimondini, 2004) by considering psychiatric in addition to medical content categories. It

https://doi.org/10.1017/S0033291709990730 Published online by Cambridge University Press

possesses satisfactory inter-rater reliability, with percentage agreements between 80% and 86% and Cohen's κ between 0.88 and 0.87 (Rimondini *et al.* 2006).

The VR-PICS consists of 39 mutually exclusive categories, 23 for psychiatrist and 16 for patient speech. For the present study only psychiatrist talk was considered, and the coded speech units were divided into three groups: patient-centred expressions; doctor-centred expressions; and neutral expressions.

Patient-centred expressions. This group contains the selected patient-centred target expressions expected to increase as a result of the training. They are expressions of active listening such as reflections and open-ended and patient agenda completing questions ('… something else?'), orienting transitions ('I need more information on this topic before going on …'), sign-posting expressions of clarifying ('what do you mean by occasional drinker?'), checking and summarizing, expressions of seeking and handling emotions ('How do you feel about this? You must feel quite upset for this news …') and asking patients' point of view.

Doctor-centred expressions. This group comprises the set of doctor-centred interventions expected to decrease after training: brief answers ('yes'/'no'), instructions and suggestions (e.g. 'you should be more self-confident'), disapproval (e.g. 'this was not a good thing to do') and closed-ended questions.

Neutral expressions. The remaining non-target expressions are included in the category 'neutral'. They include information giving, simple reassurance ('Don't worry!'), and undemanding behaviour such as simple facilitations ('Hmm, Ah …') and transitions ('Now another question …'), conversation, asking for repetition ('Sorry?') or for understanding ('Okay?') and agreements ('You're right!').

Measure of interview performance

For each consultation a performance index was derived, defined by the ratio of the selected patient-centred to doctor-centred expressions (PtC:DcC) used by the resident. This index was used as an outcome variable to assess training effectiveness. The index increases with increasing use of patient-centred skills and/or with decreasing doctor-centred expressions.

Design

ITS design explores the performance trend within a series of observations and checks if the trend observed is modified by an intervention. Data are collected at multiple instances over time before and after the intervention (here training) to detect whether the intervention has an effect significantly greater than the underlying initially observed trend (Ramsey *et al.* 2003). To increase the goodness of fit of the baseline trend, and consequently the validity of the training effect estimation, a greater number of pre-training than post-training interviews were collected.

Statistical analysis

Unpaired t tests were performed to check whether there was a difference between the mean frequency of the coded speech units per interview before and after the training.

Kolmogorov–Smirnov two-sample tests (D) assessed the effect of the confounding variables [patient gender and diagnosis (anxiety *versus* depression)] on the performance index PtC:DcC at T1.

The individual pre-training performance indices allowed the determination of the best applicable baseline regression model. First, visual inspection of each resident's time series confirmed a linear form of trends. Second, to explore the presence of heterogeneity among residents' series, within- and between-resident variances were calculated and a linear robust regression estimate (specifying residents' dummy variables and order of interviews as explanatory variables) was determined to obtain the specific level of each resident (Greene, 2003). As a third step the presence of first-order autocorrelation was tested using the Wooldridge test for panel data (Drukker, 2003). Finally, a cross-sectional time-series linear model with first-order autoregressive errors was fitted to check the presence of 'natural' trends of improvement (here practice effect).

After that, a multilevel segmented regression analysis of ITS on panel data was performed to estimate the following equation:

$$y_{ij} = a_0 + b_0 t_{ij} + [a_1 + b_1(t_{ij} - c)]d + e_{ij} + u_j,$$

where y_{ij} is the performance index in the ith interview of the jth resident, t_{ij} is the order of interviews (1 to 12) of the jth resident, $a_0 + b_0 t_{ij}$ is the linear baseline component estimated before the training (in detail: a_0 is the intercept, which expresses the constant level; b_0 is the slope, which expresses the 'natural' trend of outcome), $a_1 + b_1(t_{ij} - c)$ is the linear variation component due to training [like above: a_1 is the intercept, which expresses the level change due to training; b_1 is the variation of slope attributable to training; c is the number of interviews at pre-training (8)], d is a dummy variable representing the phase of study (pre-post training), and $e_{ij} + u_j$ is the error term at time i for the resident j (representing the random variability not

https://doi.org/10.1017/S0033291709990730 Published online by Cambridge University Press

Table 1. *Distribution of patient- and doctor-centred expressions per interview (expressed by VR-PICS categories) comprising the performance index*

	Expressions, frequencies per interview	
	Pre-training	Post-training
Patient-centred skills		
Information gathering		
Reflecting	8.4 (6.8)	9.2 (6.5)
Open question	6.5 (3.7)	5.3 (2.8)
Explaining transition	2.0 (1.8)	4.3 (2.5)
Clarifying, checking, summarizing	10.2 (6.7)	11.4 (7.6)
Asking for opinion	4.7 (3.9)	3.9 (2.4)
Completing agenda	1.2 (1.4)	2.1 (1.7)
Total	33.1 (13.9)	36.2 (15.6)
Handling emotion		
Appraisal (legitimizing, empathy, participation, respect)	1.3 (1.3)	3.6 (2.0)
Open question on emotion	2.3 (1.9)	2.4 (2.3)
Total	3.5 (2.0)	6.0 (3.0)
Total	36.6 (14.3)	42.2 (17.0)
Doctor-centred expressions		
Giving instructions	0.3 (0.6)	0.1 (0.2)
Closed question	26.7 (13.3)	17.7 (11.1)
Brief answer	0.4 (0.7)	0.5 (1.0)
Disapproval	0.0 (0.1)	0.0 (0.2)
Total	27.4 (13.3)	18.3 (11.1)
Neutral expressions		
Total	42.6 (19.4)	36.8 (19.2)

VR-PICS, Verona Psychiatric Interview Classification System.
Values are given as mean (standard deviation).

explained by the model, and can be divided in two components: one due to the *i*th interview within the *j*th resident, the other to the *j*th resident).

This equation assesses how much the training changed the interview performance, immediately and over time, producing different kinds of change, gradual (b_1) or abrupt (a_1) (Wagner *et al*. 2002). All 10 performance profiles were assessed simultaneously by a cross-sectional time-series analysis with a first-order autocorrelation. Fixed-effect estimators were chosen to take into account the different characteristics of the residents rather than their representativeness.

To assess the variability among and within residents, analysis of variance (ANOVA) and the intra-class correlation coefficient (ICC) were calculated for both pre- and post-training performance indices. The analyses were performed using Stata version 9.2 (StataCorp LP, USA).[1]†

Results

The analyses were based on 118 consultations (two pre-training videotapes of resident no. 9 were damaged), comprising a total of 12204 coded verbal speech units of residents' talk. There was a mean of 103 speech units per interview (S.D. = 33.2). No difference emerged in the mean speech units per interview at T1 and T2 [106.6 (S.D. = 31.4) and 97.3 (S.D. = 36.2), respectively, $t = 1.39$, $p = 0.17$]. The performance index PtC:DcC derived from the pre-training interviews was unaffected by the potentially confounding variables of SP gender (1.8 and 1.5 for female and male gender, respectively; D = 0.19, $p = 0.38$) and psychopathology (1.8 and 1.5 for anxiety and mood disorder, respectively; D = 0.21, $p = 0.28$).

Table 1 shows the overall distribution of patient-centred and doctor-centred expressions (numerator and denominator, respectively, of the performance index) at pre- and post-training. The observed increase of patient-centred expressions, particularly emotion handling skills, and decrease of doctor-centred

† The notes appear after the main text.

https://doi.org/10.1017/S0033291709990730 Published online by Cambridge University Press

Table 2. *Mean frequencies of patient-centred skills (information gathering and emotion handling) and doctor-centred expressions per interview and mean performance index at pre- and post-training, by resident*

	Information gathering		Handling emotion		Doctor-centred		Performance index	
Training …	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Residents								
1	30.7	26.5	3.4	3.5	40.8	25.5	0.9	1.7
2	43.4	60.3	3.1	7.5	27.3	23.3	1.8	4.2
3	39.5	27.5	4.5	6.8	40.3	16.5	1.2	2.1
4	37.8	45.5	3.4	4.8	21.5	21.5	2.1	2.4
5	37.1	56.3	3.8	9.0	17.9	26.3	3.1	3.0
6	24.5	23.3	3.8	4.5	20.0	16.0	1.5	1.8
7	38.0	34.5	2.6	3.8	33.4	12.5	1.2	3.2
8	21.7	26.3	3.8	3.5	35.8	20.3	0.7	1.5
9	28.0	28.8	2.8	7.5	22.8	8.3	1.3	4.7
10	28.8	33.8	3.9	9.0	13.4	13.3	2.5	5.7
Total	33.9	36.3	3.5	6.0	27.4	18.3	1.6	3.0

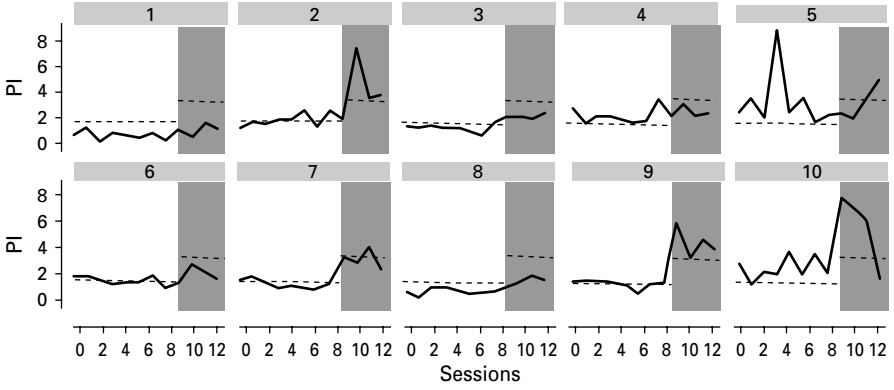

Fig. 2. Line plot of performance index (PI; ratio of the patient-centred to doctor-centred expressions) per interview by residents (1–10) before (sessions 1–8) and after training (sessions 9–12). - - -, Regression slope.

expressions, particularly closed questions, as well as the overall increase of the performance index from 1.6 (S.D. = 1.1) to 3 (S.D. = 1.9) at post-training (Table 2) suggests the sensitivity of the target index to hypothesis-expected changes. Table 2 shows how residents vary in their performance and performance change. The heterogeneity assumption is confirmed by ANOVA (pre-training: $F = 5.72$, $p < 0.001$; post-training: $F = 4.44$, $p < 0.001$) and ICC values of 0.34 [confidence interval (CI) 0.13–0.62] for pre- and 0.43 (CI 0.15–0.75) for post-training. Robust regression on pre-training performance indices showed relevant differences among residents (residents no. 1, 3, 6, 7, 8, 9 *versus* residents no. 2, 4, 5, 10) and the absence of a practice effect (estimated slope = −0.01, $p = 0.80$). The Wooldridge test on pre-training series confirmed the assumption of autocorrelation [$F(1, 9) = 85.4$, $p > 0.01$].

The regression model evidenced a horizontal slope at T1 and T2 [$b_0 = -0.02$ (S.D. = 0.08), $p = 0.78$; $b_1 = -0.12$ (S.D. = 0.19), $p = 0.53$], suggesting the absence of a significant effect due to the course of time. The significant level change indicates a positive training effect [$a_1 = 1.75$ (S.D. 0.55), $p < 0.01$]. The between- and within-resident variability increased after the training (S.D. from 0.74 to 1.47 and from 0.83 to 1.22, respectively).

Fig. 2 shows for each resident the individual performance profile based on his/her 12 consultations, with the estimated pre- and post-training regression slopes. Two residents (no. 5 and no. 10) showed an irregular performance over all pre-training interviews that was maintained at T2. Two residents (no. 2 and no. 7) showed a particularly irregular profile at post-training.

Discussion

This study evaluated the effects of a communication skills training for psychiatric residents, in producing measurable changes in interviewing behaviour, defined by a performance index based on observed

https://doi.org/10.1017/S0033291709990730 Published online by Cambridge University Press

frequencies of a selected set of desirable and undesirable behaviours in a series of interviews with unannounced SPs. Our analyses demonstrated a significant performance change in the four interview sessions after the training and allowed us to exclude a practice effect on residents' interview performance in the eight pre-training sessions. Training tended to increase desirable skills, particularly emotion handling, and decreased undesirable verbal behaviour, particularly closed-ended questions.

The uneven performance of residents over their four post-training interviews, compared with that at pre-training, and the corresponding sudden rise of the variability of the performance index was unexpected and suggested difficulties to put the learned skills into practice.

Strengths

The evaluation of the effect of training courses on interview performance in psychiatric settings does not easily lend itself to the collection of data. Differences in consultation length, patients' sociodemographic and clinical characteristics, physicians' attitude and personal style have been shown to affect the interviewing performance of physicians during consultations (Hall & Roter, 1995; Levinson & Roter, 1995; Del Piccolo *et al*. 2000; Freeman *et al*. 2002) and are likely to interact with training effects also in psychiatry. To account for this variability in a randomized clinical trial design would have demanded sample sizes of residents, patients and consultations beyond the available economic and human resources. The ITS design offered an acceptable alternative. To our knowledge this is the first time that the assessment of interview performance of health professionals was based on repeated interview sessions before as well as after the training and that an ITS design was used in evaluating the interview performance over time. Moreover, the individual training effect of each participant was taken into account in the application of multilevel analysis.

The study met the Cochrane Group criteria of an ITS design (Bero *et al*. 2002): at least three time points before and after the intervention; the intervention occurring at a clearly defined point in time; and health provider performance measured objectively.

The balanced use of 20 simulated patients, each portraying a different, although standardized script, assured greater homogeneity in terms of psychosocial data and number and type of symptoms and guaranteed that the interview tasks were similar between and within residents and not affected by SP gender and the presented psychopathology. Residents' blindness was maintained until the end of the study, facilitated by the realistic setting in which SPs were interviewed. It could be argued that some residents may have suspected that patients were not real and that this could have altered their performance. However, Maguire *et al*. (1986) demonstrated that although doctors were told that some of their patients were simulators they could not tell who these were.

Limits

The 10 participants represented the whole population of second-year residents in psychiatry at the University of Verona in 2002, and their interview performance, here expressed as an index, may not be representative of psychiatric residents in general. However, this measure proved to be sensitive to changes determined by the training course and as such could be utilized in other population samples of health providers independently of their performance level at pre-training.

The changes in the performance index after training do not inform us if they were due to an increase in patient-centred techniques, a decrease in doctor-centred interventions, or to both. Due to the small sample size the observed pre-/post-training differences could not be tested for statistical significance, but trends suggested that changes were particularly pronounced for emotion handling and closed enquiry. The high instability of our outcome measure over the four post-training observations would have required a greater number of observations and increased the robustness of our results. Given the new frontier nature of our study we could not anticipate the destabilizing effect of the training and the difficulty of residents to put their learning into practice. Without any preceding evidence on this phenomenon in the literature we privileged the pre-training performance estimates of each resident. Moreover, given the lack in the literature on repeated observation of communication skills over time, power calculations could not be performed.

Main findings

No gold standard can exist defining what would be the appropriate frequencies of single doctor- or patient-centred interventions in psychiatric consultations. However, inspection of the mean use of some patient-centred and doctor-centred expressions suggests that pre-training interviews were characterized by many closed-compared with open-ended questions, a modest use of orienting transitions and by rare emotion handling expressions (an interviewing approach examined in detail in a first analysis of this dataset; Rimondini *et al*. 2006). Similar consultation characteristics have been previously observed in psychiatry (Maguire, 1982; Goldberg *et al*. 1984) and in

other medical specialties, such as general medicine and gynaecology (Suchman *et al*. 1997; Van Dulmen *et al*. 2003; Del Piccolo *et al*. 2005). Psychiatrists were reported to be more empathic than other physicians (Hojat *et al*. 2002) and to have better relational skills (Sierles *et al*. 2004). These observations contrast with the infrequent use of emotion-handling skills of our residents at pre-training, despite the longer duration of the psychiatric compared with other medical consultations, and the greater prevalence of psychosocial and emotional problems, competently simulated by our SPs.

The increase of emotion-handling interventions after the training represents an affirmative answer to one debated issue in clinical education: whether empathy and acknowledging of patient emotions can be taught. The finding, however, refers to verbal expressions. Non-verbal indicators, crucial to modulate the verbalized emphatic message, were not considered here.

The reduction in the frequency of closed-ended questions was another desirable course effect, but was not substituted, as expected, by a parallel increase in the use of open questions. This last learning target was not achieved and should be considered in future training courses, although the observed increase in reflections, transitions and other patient-centred interventions at post-training could have diminished the need of directive open-ended questioning. An alternative hypothesis is that training improved the quality of open-ended questions so that fewer questions were needed to obtain the required information. Additional qualitative analyses that would complete the picture and allow a deeper understanding of residents' improvements were beyond the scope of this paper.

The interview performance index varied between and within residents before and particularly so after training. These findings have implications for communication research and teaching. The performance variation of residents, expressed by the ICC at pre-training, will help future studies in this field perform the recommended power calculations for the determination of sample size and expected size of improvement in performance (Bland, 2000). The observed differences between residents before training parallel previous findings that course participants rarely start from the same level of expertise and have a distinctive interviewing style (Hulsman *et al*. 2005). Teaching should be tailored to individual learning needs, to be assessed beforehand. Different approaches and skill levels may initially generate conflict in trainees, but become powerful teaching resources by offering the chance to share opinions and to learn from peers (Lieberman *et al*. 1978). The residents' performance profile as well as the pre-/post-training comparison in the use of patient-centred expressions and closed inquiry confirmed that individuals have different ways of putting into practice what has been learnt (Lieberman *et al*. 1989; Harrison & Goldberg, 1993; Nuzzarello & Birndorf, 2004). Moreover, the increase in the performance variability within residents from pre- to post-training implies that assessment of interviewing skills based on a single interview would not present a realistic estimate of the learner's performance. Post-training exercise and supervised practice are needed in order to establish a more regular performance at a higher skill level which in order to become observable would require a prolonged follow-up period for evaluation.

In conclusion, our findings on training effects are encouraging concerning the acquisition of skills, but their full integration into practice has still to be demonstrated, suggesting that special thought has to be dedicated to the organizational and content-related aspects of the follow-up period. The ITS design described here appears to be a useful method for the evaluation of time-limited interventions in psychiatry. Future research may take advantage of the parameters and observations offered in order to confirm and extend the results of the present study.

Acknowledgements

This paper received no external funding. The authors thank Professor G. Dunn of the University of Manchester (UK) for statistical advice.

Declaration of Interest

None.

Notes

[1] The assumptions of time series were explored by the following Stata commands: 'xtsum', 'loneway',' xtserial' and 'regression' with dummy variables for residents and option robust to explore the level of each resident. Panel regression was estimated using the 'xtregar' command with option fe.

References

Bero L, Grilli R, Grimshaw JM, Oxman A (2002). The Cochrane Effective Practice and Organisation of Care Group (EPOC) Module. In *The Cochrane Library, Issue 1*. Update Software: Oxford, UK.

Bland JM (2000). Sample size in guidelines trials. *Family Practice* **17**, 17–20.

Bodenheimer T, Lorig K, Holman H, Grumbach K (2002). Patient self-management of chronic disease in primary

care. *Journal of the American Medical Association* **288**, 2469–2475.

Butow P, Dunn S, Tattersall M, Jones Q (1995). Computer-based interaction analyses of the cancer consultation. *British Journal of Cancer* **71**, 1115–1121.

Campbell DT, Stanley JC (1966). *Experimental and Quasi-Experimental Designs for Research*. Houghton Mifflin Co.: Boston, MA.

Del Piccolo L, Mead N, Gask L, Mazzi M, Goss C, Rimondini M, Zimmermann C (2005). The English version of the Verona medical interview classification system (VR-MICS). An assessment of its reliability and a comparative cross-cultural test of its validity. *Patient Education and Counseling* **58**, 252–264.

Del Piccolo L, Putnam S, Mazzi MA, Zimmermann C (2004). The biopsychosocial domains and the functions of the medical interview in primary care: construct validity of the Verona Medical Interview Classification System. *Patient Education and Counseling* **53**, 47–56.

Del Piccolo L, Saltini A, Zimmermann C, Dunn G (2000). Differences in verbal behaviours of patients with and without emotional distress during primary care consultations. *Psychological Medicine* **30**, 629–643.

Drukker DM (2003). Testing for serial correlation in linear panel-data models. *Stata Journal* **3**, 168–177.

Freeman GK, Horder JP, Howie JG, Hungin AP, Hill AP, Shah NC, Wilson A (2002). Evolving general practice consultation in Britain: issues of length and context. *British Medical Journal* **324**, 880–882.

Goldberg DP, Hobson RF, Maguire GP, Margison FR, O'Dowd T, Osborn M, Moss S (1984). The clarification and assessment of a method of psychotherapy. *British Journal of Psychiatry* **144**, 567–575.

Goss C, Del Piccolo L, Mazzi MA, Moretti F, Rimondini M, Zimmermann M (2008). Involving patients in decisions during psychiatric consultations. *British Journal of Psychiatry* **193**, 416–421.

Greene WH (2003). *Econometric Analysis*. Prentice Hall: New Jersey.

Hall JA, Roter DL (1995). Patient gender and communication with physicians: results of a community-based study. *Women's Health* **1**, 77–95.

Harrison J, Goldberg D (1993). Improving the interview skills of psychiatry trainees. *European Journal of Psychiatry* **7**, 31–40.

Hojat M, Gonnella JS, Nasca TJ, Mangione S, Vergare M, Magee M (2002). Physician empathy: definition, components, measurement, and relationship to gender and specialty. *American Journal of Psychiatry* **159**, 1563–1569.

Hulsman R, Visser A, Makoul G (2005). Addressing some of the key questions about communication in healthcare. *Patient Education and Counseling* **58**, 221–224.

Joyner B, Young L (2006). Teaching medical students using role play: twelve tips for successful role plays. *Medical Teacher* **28**, 225–229.

Kurtz S, Silverman J, Draper J (1998). *Teaching and Learning Communication Skills in Medicine*. Radcliff Medical Press: Abingdon, UK.

Levinson W, Roter D (1995). Physicians' psychosocial beliefs correlate with their patient communication skills. *Journal of General Internal Medicine* **10**, 375–379.

Lieberman S, Cobb J, Jackson CH (1989). Studying the 'Grammar of Psychotherapy' Course using a student and control population. Some results, trends and disappointments. *British Journal of Psychiatry* **155**, 842–845.

Lieberman S, Hafner RJ, Crisp A (1978). Teaching psychotherapy in mental hospitals. *British Journal of Psychiatry* **132**, 398–402.

Maguire GP, Goldberg DP, Hobson RF, Margison F, Moss S, O'Dowd T (1984). Evaluating the teaching of a method of psychotherapy. *British Journal of Psychiatry* **144**, 575–580.

Maguire P (1982). Psychiatrists also need interview training. *British Journal of Psychiatry* **141**, 423–424.

Maguire P, Fairbairn S, Fletcher C (1986). Consultation skills of young doctors: benefits of feedback training in interviewing as students persist. *British Medical Journal* (*Clinical Research Edition*) **292**, 1573–1576.

Morgan DG (1999). 'Please see and advise': a qualitative study of patients' experiences of psychiatric outpatient care. *Social Psychiatry and Psychiatric Epidemiology* **34**, 442–450.

Nosé M, Barbui C, Tansella M (2003). How often do patients with psychosis fail to adhere with treatment programs? A systematic review. *Psychological Medicine* **33**, 1149–1160.

Nuzzarello A, Birndorf C (2004). An interviewing course for a psychiatry clerkship. *Academic Psychiatry* **28**, 66–70.

Percudani M, Belloni G, Contini A, Barbui C (2002). Monitoring community psychiatric services in Italy: differences between patients who leave care and those who stay in treatment. *British Journal of Psychiatry* **180**, 254–259.

Ramsey C, Matowe L, Grilli R, Grimshaw JM, Thomas RE (2003). Interrupted time series designs in health technology assessment: lessons from two systematic reviews of behavior change strategies. *International Journal of Technological Assessment in Health Care* **19**, 613–623.

Rimondini M (2004). The psychiatric interview: evaluation of a training in communication skills. Ph.D. Thesis, University of Verona.

Rimondini M, Goss C, Del Piccolo L, Mazzi M, Paccaloni M, Zimmermann C (2006). Communication skills in psychiatry residents: how do they handle patient concerns? An application of sequence analysis to interviews with simulated patients. *Psychotherapy and Psychosomatics* **75**, 161–169.

Rossi A, Amaddeo F, Ruggeri M, Thornicroft G, Tansella M (2002). Dropping out of care. Inappropriate terminations of contact with community based psychiatric services. *British Journal of Psychiatry* **181**, 331–338.

Sierles FS, Vergare MJ, Hojat M, Gonnella JS (2004). Academic performance of psychiatrists compared to other specialists before, during, and after medical school. *American Journal of Psychiatry* **161**, 1477–1482.

Stewart M, Brown JB, Donner A, McWhinney IR, Oates J, Weston WW, Jordan J (2000). The impact of patient-centered care on outcomes. *Journal of Family Practice* **49**, 796–804.

Suchman AL, Markakis K, Beckman HB, Frankel R (1997). A model of empathic communication in the medical interview. *Journal of the American Medical Association* **277**, 678–682.

Swick S, Hall S, Beresin E (2006). Assessing the ACGME Competencies in Psychiatry Training programs. *Academic Psychiatry* **30**, 330–351.

Van Dulmen S, Nuebling M, Langewitz W (2003). Doctors' responses to patients' concerns; an exploration of communication sequences in gynecology. *Epidemiologia e Psichiatria Sociale* **12**, 98–102.

Wagner AK, Soumerai SB, Zhang F, Ross-Degnan D (2002). Segmented regression analysis of interrupted time series studies in medication use research. *Journal of Clinical Pharmacy and Therapeutics* **27**, 299–309.

Walton H, Gelder M (1999). Core curriculum in psychiatry for medical students. *Medical Education* **33**, 204–211.

Young A, Grusky O, Jordan D, Belin T (2002). Routine outcome monitoring in a public mental health system: the impact of patients who leave care. *Psychiatric Services* **51**, 85–91.