

Supplemental Digital Content 2. Main features of the studies included in the systematic review

STUDY, LOCATION	DESIGN	PARTICIPANTS	TYPE OF INTERVENTION	MAIN OUTCOMES	MAIN FINDINGS	MESQRI (/18)	Kirkpatrick level
Bosse et al, 2010 Germany	RCT Two groups: 1) PRP group 2) SimP group	5th year N= 69 Age = NA Sex ratio = NA Response rate : 96.7%	Theme: communication with parents of sick children Content : training sessions including : - seminars - problem-based learning - virtual patients (CAMPUS) - bedside teaching - SimP or PRP training cases SimP and PRP training cases were chosen using focus group with teachers. Each training session include 3 training cases Total duration : 4 weeks (3 sessions, 2h30 each) Group size : 3 students per group	Satisfaction assessed using a questionnaire, which included 5 dimensions: acceptability, realism, interest, perceived usefulness of situations, further applicability. Instrument without published validity data	No between group difference for acceptability, realism, and perceived usefulness of situations Statistically more “interest” and “further applicability” for the SimP program according to students	11	Level 1
Bosse et al, 2012 & 2015 Germany	RCT Three groups: 1) PRP 2) SimP group 3) “no education”	5th year N= 103 Age = NA Sex ratio = NA Response rate : 88.2 to 100%	Theme: communication with parents of sick children Content : training sessions including : - seminars - problem-based learning - virtual patients (CAMPUS) - bedside teaching - SimP or RP training cases SimP and PRP training cases were chosen using focus group with teachers. Each training session included 3 training cases Total duration : 4 weeks (3 sessions, 2h30 each) Group size : 3 students per group	Self-efficacy on communication skills: 24-item questionnaire (14 items addressed general competences, and 10 items addressed cases-specific learning objectives), with 10 points likert rating. Total score range 0-10. Instrument without published validity data Objective Structured clinical examination, comprising 6 stations with SimP. Evaluation with a standardized checklist (Calgary-Cambridge Referenced Observation Guide), by trained psychologists Instrument with published validity data.	Self efficacy Pre-test : NS between-group difference Post-test : - SimP > control group (p=0.007, Cohen’s d 0.32) - PRP group > control group (p= 0.021, Cohen’s d 0.673). - No difference between SimP and PRP group (p=0.704) Communication Skills Post-test : - PRP group > “no education” group (p=0.006, Cohen’s d 1.48) - SimP group > control group (p=0.001, Cohen’s d= 0.63). - PRP group > SimP group (p=0.021, Cohen’s d = 0.71). - subscore analyses: statistically significant difference for « understanding parent’s perspective » in the PRP group, compared to the SimP group.	14	Level 2b Level 2d

				<p>Cost-effectiveness analysis</p> <p>Effectiveness : Main score using the Calgary Cambridge Referenced Observation Guide (score in %)</p> <p>Costs :</p> <ul style="list-style-type: none"> - SimP Training - SimP employment for roleplay - SimP training for SimP group - Tutor costs <p>Ratio of effectiveness : (OCSE score in %) divided by man-hours (in h) for both groups</p>	<p>Cost-effectiveness analysis:</p> <p>Effectiveness :</p> <ul style="list-style-type: none"> - PRP group : 81.6% ±3.32 - SimP group : 78.0% ± 6.23 <p>Costs :</p> <ul style="list-style-type: none"> - SimP training:: 5 man-hours - SimP employment for roleplay: 21 man-hours - SimP training for SimP group: 60 man-hours - Tutor costs : 86 man-hours in both group <p>Ratio of effectiveness :</p> <ul style="list-style-type: none"> - PRP group = 0.74 - SimP group = 0.45 		
Coonar et al. 1991 United Kingdom	Uncontrolled study Post-test only	1 st year N = 21 Age = NA Sex ratio = NA Response rate : NA	<p>Theme: PRP in teaching communication skills in medical students during gynecology examination.</p> <p>Content: 3 PRP case studies (couple infertility, positive cervical smear test and couple requiring sterilization of the female partner). A 3-minute "time out" could be called at any time by anyone ("doctor", "patient" or "audience") during the PRP sessions.</p> <p>Total duration: 2.5h (3 sessions of 40 minutes + general debriefing)</p>	Post-test questionnaire with 4 items related to acceptability, interest, perceived transferability into practices. Instrument without published validity data	PRP thought to be an effective technique to teach a conventional specialty	6	Level 1
Cushing et al, 1995 United Kingdom	Uncontrolled study Pre and post tests	4 th and 5 th year N= 231 Age = NA Sex ratio = NA Response rate : 65%	<p>Theme: evaluation of breaking bad news course</p> <p>Content : group discussion, videotape presentation, PRP. Students developed their own scenarios, defined their own learning objectives. Feedback was based on the roleplay records on a videotape. 3 groups of medical students were:</p> <ul style="list-style-type: none"> - telling parents about their newborn child - telling someone about a close relative - telling someone about himself <p>Total duration : 6 hours (2 sessions of 3 hours) Group size : 6-10 students per group</p>	<p>Satisfaction and knowledge assessment: students complete a « before and after » evaluation form about level of confidence in breaking bad news, their knowledge, their experience of the course.</p> <p>Instruments without published validity data</p>	<p>Satisfaction: High satisfaction ratings, with 82 to 97% of participants who recommended PRP, depending on groups.</p> <p>Knowledge : Significant improvement of knowledge (particularly in relation to interpersonal communication), more confident No statistically significant difference between 4th and 5th years groups</p>	10	Level 1 Level 2a Level 2d
Fertleman et al 2005 United Kingdom	Uncontrolled study Post-test only	2 nd year N= 205 Age = NA Sex ratio = NA Response rate : 80%	<p>Theme : communication skills with parents of ill children</p> <p>Content :</p> <ul style="list-style-type: none"> - 30 minutes video of mother of ill children talking about their (bad) experiences, - 5 minutes RP with one of mother's situation and 5 minutes feed-back - 20 minutes discussion <p>Total duration : 1h Group size : 12</p>	<p>Satisfaction assessment</p> <p>Students scored usefulness session evaluation using not detailed questionnaire for video, RP and discussion (0 to 100 points Likert scale). Instruments without published validity data</p>	<p>Satisfaction : PRP felt to be useful (Likert scores videos 84, RP 68, discussion 77, whole session 73)</p>	8	Level 1

Gartmeier et al 2014 Germany	RCT Four groups: 1) PRP group 2) e learning group 3) PRP + e learning group 4) "no education" group	3 rd year N = 72 Age = NA Sex ratio = 0.32	Subject: Investigated the effectiveness of three version of a training program on physician-patient and teacher-parent conversations for medical students and student teachers Content: Analysis of the effects of the different applied training methods and determine the similarities across professional domains. Video cases, PRP + video feedback. RCT of 3 conditions: - e-learning analyzing video cases of professional conversations - RP and video feedback - their combination - group control with no training Total duration: 300 min Group size: 16 (4 per condition)	Communication skills assessment: Videotaped SP, evaluated by 2 trained raters analyzing 3 competence facets according Seidel and Walker and Dogter with a 9 items rated on five-point Likert scale. Instrument with published validity data.	Communication skills : - Strong overall treatment effect versus control group (p<0.001) - Combined condition = more effective than e-learning and RP alone (p<0.02) - e-learning proved more effective than RP (p=0.007)	13	Level 2b
Kiosses et al., 2017 [18] Greece	RCT Two groups: 1) PRP group 2) "no education" group	4 th to 6 th years N=87 Age= NA Sex ratio= 0.51	Theme: Training in empathy Content: Theory, personal development and skills development: The skills development section contained active listening exercises, role plays, non-verbal communication games, case studies, conduction of a person-centered interview and practical implications of empathy Total duration: 3 20-hour workshops four weeks apart from each other Group size: NA	Jefferson Scale of Physician Empathy (JSPE): self-questionnaire (20 items, 7 points Likert scale). Evaluation before, after course and 6 months after. . Instrument with published validity data.	Empathy skills : - PRP> control group after training (p<0.001, Cohen's d = 1.268) and at 6 month (p<0.001, Cohen's d = 1.238)	10	Level 2b
King et al., 2015 Australia	Uncontrolled study post-test only	5 th and 6 th years N=107 Age= 21 to 30 years Sex ratio: NA Response rate:79%	Theme: psychiatry (anxiety disorders, eating disorders, psychotic disorders ...) Content: role-play based learning, with students playing both doctor and patient. Feedback discussion with senior psychiatrist Total duration: 1 to 6 sessions Group size: NA	Participation : number of sessions attended Satisfaction, by questionnaire (12 items, 5 points Likert-scale). Instrument without published validity data	Participation : average number of session : 5.4 Satisfaction: Respondents were positive about all aspects of the RP Based Learning sessions (Median =4.1±0.3)	8	Level 1
Koponen et al., 2012 Finland	RCT Three groups: 1) PRP 2) SimP group 3) theater group	2 nd years N=129 Age=22.2 years Sex ratio=0.51	Theme: communication skills in general practice Content: - Lesson in physician– patient communication and analysis of physician–patient interaction shown on DVD. - Practicing communication skills randomly through the theatre, SimP or PRP group. - Observing and analyzing real doctor–patient encounters (2 days in health centers). - Discussing learning experiences in small groups. Total duration: 3 months Group size: 6 (SimP or PRP group) to 16 (theater group)	Communication skills attitude using the CSAS Scale (Communication Skills Attitude Scale): Instrument with published validity data., which measures medical students' attitudes towards communication skills learning (26 items, five-point Likert scale. 13 items are related to negative attitudes and 13 to positive attitudes)	Communication attitude: all group statistically enhanced their skills, without difference between groups. -Female students had more positive attitudes than male students (Cohen's d = NA).	13	Level 2c
Koponen et al., 2014	RCT Three groups:	2 nd years N=132 Age=22. years	Identical to Koponen et al, 2012 - Theater group = 14 to 16 students, 2 facilitators and a theatre cast. Actors played a	Communication attitude using the CSAS Scale (see above)	- Communication attitude: all group statistically enhanced their skills, without difference between groups.	12	Level 2c

Finland	1) PRP 2) SimP group 3) theater group	Sex ratio =0.67	performance included communication challenge (not making eye contact, not responding to patient's emotions etc). Feed-back discussions. - SimP group : 6 students, a facilitator. Medical students played the doctor role, with SP. Feed-back discussions. - PRP group : 6 students, a facilitator. Medical students played doctor's and patient's roles. Feed-back discussions.		-Female students had more positive attitudes than male students (Cohen's d = NA).		
Lau et al 2001 Hong Kong	RCT cross over design Two groups: 1) observer group 2) interpreter group	1st year N=160 Age : NA Sex ratio : NA	Theme : general communication skills Content : 3-part course : - video of patient interview - RP - Feedback from tutors Students played RP all roles in both groups, as follow : - Observer group : doctor, patient (Cantonese or English speaking), and observer - Interpreter group : doctor, patient (Cantonese speaking) and interpreter (bilingual, accompanying the patient) Total duration : 4h30 Group size : 3	Communication skills knowledge Self-questionnaire, 9 items 5 points Likert. Instrument with published validity data Perceived effectiveness of training program Self questionnaire, 7 items 5 points likert. Instrument with published validity data	Communication skills knowledge: Observer group statistically enhanced communication skills compared to interpreter group (p=0.038, Cohen's d=0.318) Perceived effectiveness: statistically significant difference in pre and post-tests, without difference between groups	11.5	Level 1 Level 2a
Lavanya et al 2016 India	Uncontrolled study post-test only	2nd year N = 96 Age = NA Sex ratio = 0.78 Response rate : NA	Theme: medication communication skills Content: - Lecture class on general aspects of communication - 6 training cases (structured script) for role play sessions Total duration: NA. At least 4 training cases to be evaluated Group size: NA a role play session required a minimum of three students (physician, patient, observer)	Satisfaction and perceived benefits on medication communication skills: questionnaire (number of items not specified, 5-point likert scoring) to explore students' perspectives on PRP as a teaching tool. Instrument without published validity data	General perceptions about role-play: innovative/worthwhile taking part in them (90%) Perceived benefits of medication communication training sessions: - Helped to acquire doctor-patient interaction skills (>80%) - Gained confidence Better retention of pharmacology concepts - Feedback appreciated	7	Level 1
Luttenberger et al. 2014 Germany	Uncontrolled study post-test only	2nd years N = 173 Age = 22 Sex ratio = 1.5 Response rate : NA	Theme: general communication skills Content: - first class: short introduction to the concept of the course, followed by a review of communication theories - next five lessons: students performed/witnessed a total of 10 role plays (anamnesis, shared decision making, prevention/motivation, breaking bad news, and "difficult" interactions) Total duration: 6 days, 2 hours on each day Group size: about 15 persons	Satisfaction : self-questionnaire (14 items, five-point likert scale). Instrument without published validity data	Satisfaction: Course highly appreciated by the students	8	Level 1

Mills et al. 2014 United Kingdom	Uncontrolled study post-test only	1 st year N= 198 Age= NA Sex ratio= NA Response rate : 83%	Theme: communication skills in general practice Content: - theoretical lecture on effective communication skills - simulated videotaped consultation based on role-play from scenarios provided (chest pain, post-natal depression, sexually transmitted infection) - videobased feed back with all student's group Total duration: 3 weeks Group size: 10 students	Satisfaction : Self questionnaire. Five closed questions (3 with Likert scale and 2 with dichotomous responses) Instrument without published validity data	Students were highly satisfied with the teaching session 80% considered the session useful and 69% strongly agreed tutors facilitated their development.	7	Level 1
Mumtaz et al. 2016 Pakistan	Uncontrolled study post-test only	1 st and 2 nd years N= 351 Age= NA Sex ratio= NA Response rate : 70.2%	Theme: communication skills in general practice Content: Role- plays were conducted at the end of each module on topics extracted from the content covered by the students. Student prepared their own scenario one week before RP. RP during at least 7 minutes, and 5 minutes for feed-back with faculty member. Total duration: one week for each module, form preparation to feed back. Group size: 12 to 14 students	Satisfaction and acceptability Self-questionnaire, 10 items related to the students' perceptions of the usefulness of role- play for integrating the knowledge of basic sciences with clinical skills, opening new avenues of thinking, alleviating potential difficulties in communicating with patients, promotion of team work, improvement of interpersonal skills and reflection on real- life experiences Instrument without published validity data	Based on students' responses, role-play: - Alleviated potential difficulties in communicating with patients: 63% (n=223). - Promoted teamwork and interpersonal Skills: 69% (n=242) - Helped to integrate knowledge of basic sciences into clinical skills: 55% (n=193) - Increase of critical thinking: 41% (n=144). - roles created reflected real life scenarios: 27% (n=96)	6.5	Level 1
Nikendei et al. 2007 Germany	RCT PRP group Control group	3 rd years N=36 Age = 23.4 years Sex ratio = 0.28	Theme : - lower limbs doppler sonography - gastric tube insertion Content : - practice on one other (doppler) or using a mannequin (gastric tube) - for PRP group only. 3 different scenarios for each situation (doppler and gastric tube). Students played the physician, the assistant and the patient role. Total duration : NA Group Size : 3	Technical skills. Videotaped sessions, randomly rated by 2 tutors, using binary checklist. Instrument without published validity data Communication skills. Videotaped sessions, randomly rated by 2 tutors, using a scale exploring for dimensions: empathy, coherence, verbal and non-verbal communication. Instrument with published validity data	Technical Skills : Non-significant difference between groups for both situations Communication skills : RP group more effective than control, for both situation (p<0.001, Cohen's d =N/A)	14	Level 2b
Papadakis et al. 1997 California USA	RCT Two groups: 1) PRP group 2) "no education" group	1 st year N= 72 Age: NA Sex ratio: NA	Theme: Smoking-cessation consultation Content: - one-hour lecture on the health effect of smoking - one-hour lecture that identified specific steps physicians can take to help patients quit smoking - practice sessions (SPs or RP, randomly) Total duration: 2 weeks Group size: 12	Communication skills : evaluation session with a SimP. Overall rating of the students by SPs. 21 items post-session questionnaire. Instrument without published validity data Cost-minimization analysis Including costs : cost of training and intervention in program for actors, administrative costs Costs were calculated for a whole group and normalized to a virtual class of 100 students.	Communication skills. Non-Significant difference between the 2 groups in their ratings by the SPs. Cost-minimization analysis Costs for SP program : 2500\$ - SP training (200\$) - SP intervention (1800\$) - Administrative staff (500\$) Costs for RP program : 500\$ - Administrative staff (500\$) SP program is 5 times more expensive than RP program	14	Level 2b

Roman et al. 2011 USA	Uncontrolled study Pre-test/post-test design	3 rd years N=98 Age: NA Sex ratio: NA Response rate : 90%	Theme: smoking cessation consultation, motivational interviewing skills Content: - video clips comparing motivational interview with traditional medical interview - RP scenarios played by students Total duration: a single 3-hour session Group size: NA	Assessment of students' confidence and knowledge Self-administered questionnaire (6-weeks time lapse between pre- and post-clerkship tests). Questionnaire, 10 items, (-points Likert scale. Instrument without published validity data	Students' confidence and knowledge. Significant increase in knowledge and confidence for 9 of the 10 items (p<0.005).	10.5	Level 2a Level 2d
Saab et al. 2005 Lebanon	Uncontrolled study post-test only	2 nd year medical students N=150 Age: NA Sex ratio: 2,6 Response rate : 97%	Theme: basic communication skills Content: - lecture on history taking and doctor-patient relationship - brainstorming session on importance of and barriers to good communications skills - video clips (N=8) - RP with video clip's scenario - debriefing session Total duration: a single 3-hour workshop Group size: 8 to 10 students	Assessment of students' satisfaction. Self-administered post-workshop questionnaire. 6 items, 5-point Likert scale. Instrument without published validity data	Students' satisfaction. Students evaluated the sessions positively for relevancy, quality of video clips, teaching methods, and usefulness for future practice	8	Level 1
Shield et al. 2011 USA	Uncontrolled study post-test only	1 st and 2 nd year N= 286 Age: NA Sex ratio: NA Response rate : 66 to 90%	Theme : To improve communication skills with patients, families and the healthcare team Content: - clinical cases and videos - group discussions - RP - skills practice - guest patient's experiences presentations Total duration: 4 sessions of 3hours Group size: NA	Assessment of students' and faculty's' satisfaction. Self-administered post-session questionnaire. Instrument without published validity data, number of items not reported	Students' and faculty's' satisfaction. 93% of faculty and 83% of students who completed the evaluations rated the sessions as good, excellent, or exceptional. 80% of students and 96% of faculty who completed the evaluations responded that they believed that students gained knowledge that will help them care for patients	5.5	Level 1
Tayem et al. 2016 Kingdom of Bahrain	Non randomized controlled study Two groups: 1) PRP group 2) "no education" group	2 nd year medical students N=179 (133 in RP group, 4- in control group) Age: NA Sex ratio: NA Response rate : 81.2%	Theme: Cardio-vascular drug therapy communication Content: - clinical cases discussion - RP Total duration: NA Group size: NA	Assessment of students' satisfaction. Self-administered post-session questionnaire. 16 items, 5-point Likert scale from 1 (strongly disagree) to 5 (strongly agree). Focus group discussion. Instrument without published validity data	Students' satisfaction. role-play demonstrations helped develop communication skills, confidence to communicate drug-related information in a prescription, and the ability to explain the aim of drug therapy to patients. Most trainees thought that they developed skills to communicate instructions on drug use.	11	Level 1 Level 2b

				<p>Structured clinical examination. The academic achievement of both RP and control group was done with SimP, and assessed by tutors with checklist without published validity data.</p>	<p>Communication skills. RP group scored statistically higher in OPSE than control group ($p=0.016$, Cohen's $d = NA$). RP group performed better in introducing themselves to patients ($P < 0.05$ Cohen's $d = NA$), explaining patient's condition and the aim of drug therapy ($P < 0.01$), and providing instructions on drug use ($P < 0.01$ Cohen's $d = NA$).</p>		
<p>Windish et al. 2005 Maryland, USA</p>	<p>RCT Two groups: 1) PRP group 2) "no education" group</p>	<p>2nd year medical students N=120 Age: 24.4 Sex ratio:1 Response rate : 100%</p>	<p>Objective: Improve medical students' communication and clinical reasoning skills</p> <p>Content:</p> <ul style="list-style-type: none"> - PRP - self-reflection - group discussion - videotaped encounters - feedback - Final evaluation by SimP <p>Total duration: weekly 3-hour sessions for 6 weeks</p> <p>Group size: 6 students</p>	<p>Assessment of Students' communication performance. Students spent fifteen minutes in each SimP encounter. SimPs completed a 30-item interpersonal checklist rating behaviors on a 5-point scale (1 = poor; 5 = excellent) for a total possible score ranging from 30 to 150 per case. 3 subscales: data gathering, establishing rapport, patient education and counseling. Instrument with published validity data [22]</p> <p>Assessment of clinical reasoning. Students generated a problem list and differential diagnoses. They also completed the Diagnostic Thinking Inventory, a self-reported questionnaire (41 questions rated on a 6-point scale). Instrument with published validity data [22]</p>	<p>Students' communication skills. SP rated RP group more favorably in establishing rapport ($p=0.05$, Cohen's $d=1.2$). NS in the total score and other subscales.</p> <p>Students' clinical reasoning. No statistical difference in Diagnostic Thinking Inventory scores. PRP group did better for the problem listed ($p=0.05$, Cohen's $d = 1.24$) and history of present illness items ($p=0.01$, Cohen's $d=1.58$).</p> <p>Psychosocial History items listed. 65% in PRP group vs 44% in control group ($p=0.008$, Cohen's $d = NA$).</p>	13	Level 2b
<p>MESQRI : Medical Education Research Quality Instrument; NA : not available; PRP: Peer Role-Play; RCT : Randomized Clinical Trial; SimP: Simulated Patients</p>							