Talking in Primary Care Trial

Statistical Analysis Plan

Trial name:	Talking in Primary Care: communication skills e-learning for practitioners
Trial registration number:	ISRCTN18010240
Protocol title and version number:	TIP trial protocol v2 12th January 2023
SAP version number:	Version 2
SAP date:	08/12/2023

To be approved and reviewed by:

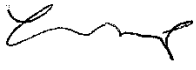
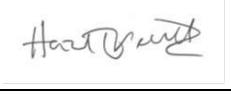
	Name	Signature	Date
Chief Investigator	Felicity Bishop		12/12/23
	Hazel Everitt		11/12/23
Statistician (author)	Taeko Becque		8/12/23
Senior Statistician	Beth Stuart		12/12/23

Table of Contents

List of Abbreviations

Abbreviation		Abbreviation	
AE	Adverse event	MAR	Missing at random
BPI	Brief pain inventory	MISS	Medical interview satisfaction scale
CARE	Consultation and relational empathy	MSK	Musculoskeletal
CI	Confidence interval	PDDOR	Patient doctor depth of relationship
CONSORT	Consolidated standards of reporting trials	PEI	Patient enablement instrument
GLMM	Generalised linear mixed model	PGI	Patient global impression
HADS	Hospital anxiety and depression score	SAP	Statistical analysis plan
IMD	Index of multiple deprivation	TEX-Q	Treatment expectation questionnaire
ITT	Intention to treat	TIP	Talking in primary care

1 Introduction

1.1 Purpose of SAP

This statistical analysis plan (SAP) describes in detail the methods that will be used to analyse the data collected as part of the TIP trial. This will form the basis of the final trial publication. The final analysis will follow the SAP to ensure that the analyses are conducted in a scientifically valid manner and to avoid post hoc decisions which may affect the interpretation of the statistical analysis. Any deviations from the SAP will be detailed in the final report.

1.2 Trial background and rationale (short synopsis)

Approximately 1.7 billion people worldwide have painful activity-limiting musculoskeletal (MSK) conditions including back, hip, knee and neck pain. Osteoarthritis is a common cause. Approximately 1 in every 7 GP consultations is about an MSK problem. Clinical guidelines recommend patient education and patient-centred care offering a range of different (ideally non-pharmacological) interventions. Regardless of which therapy a patient receives, excellent practitioner-patient communication has the potential to enhance such patient-centred care and can improve outcomes that are important to patients. Doctors communicating clinical empathy and realistic optimism about treatment outcomes can lead to reductions in pain and improvements in patient satisfaction with consultations. However, few interventions have been tested clinically for effects on patients' health, few have been sufficiently well described to allow implementation, and (where details are available) most interventions are prohibitively complex, expensive, and time-consuming. This makes engagement and uptake of current interventions extremely unlikely, particularly in the current climate of exceptionally high demand for primary care services and staff shortages. Furthermore, the COVID-19 pandemic has changed the landscape of primary care. Telephone consultations are much more widespread and present both challenges and opportunities to practitioners and patients. The pandemic has highlighted and exacerbated health inequalities and disparities in access and outcomes and emphasized the need to ensure high quality communication optimised for patients from diverse ethnic backgrounds.

We have recently developed EMPathicO, an engaging, feasible, brief, evidence-based and theoretically grounded e-learning package for primary care practitioners. Our feasibility study, and an earlier pilot trial of some EMPathicO components ('KEPE-Warm') strongly suggest that (1) EMPathicO could help practitioners enhance their communication of clinical empathy and realistic optimism and (2) EMPathicO is ready to be evaluated in a cluster-randomized trial in primary care.

Building on our successful development and feasibility work we now want to determine the clinical and cost effectiveness of EMPathicO training for primary care practitioners, primarily in patients presenting with MSK pain. We also aim to maximize the potential future impact of EMPathicO by also testing the effect of training practitioners on outcomes in patients with other conditions.

1.3 Objectives

The primary aim is to determine the clinical and cost-effectiveness of EMPathicO training in Clinical Empathy and conveying realistic Positive Messages for practitioners in patients presenting with MSK pain. The associated objectives are to:

a. Determine the effects of EMPathicO on (a) patient-reported pain and (b) patient enablement based on repeated measures over 6 months following the index consultation, in patients presenting with MSK pain.
b. Compare the costs and consequences and estimate the cost-effectiveness of EMPathicO versus usual care, over the 6 months following the index consultation, for patients with MSK pain.
c. Determine the effects of EMPathicO on patient-reported quality of life and other secondary outcomes across the 6 months following the index consultation.
d. Test the hypothesized mechanisms of action of EMPathicO, including intervention usage and effects on patient-perceived practitioner empathy and optimism (as per our logic model).

The secondary aim is to explore EMPathicO's potential for impact on conditions other than MSK pain and ways to maximise wide-spread adoption, implementation, and maintenance of effects. We will do this by assessing effects of EMPathicO training on patients presenting with any symptoms other than MSK pain since the impact of EMPathicO will potentially be in all consultations not just MSK consultations; testing how and in what circumstances EMPathicO changes practitioner communication behaviours and patient outcomes for in-person, telephone, and video consultations; and analysing a diverse range of patients' and practitioners' experiences of adoption and longer-term implementation. The associated objectives are to:

e. Determine the effects of EMPathicO on patient enablement, patient-reported quality of life and other secondary outcomes across the 6 months following the index consultation, in patients presenting with symptoms other than MSK pain.
f. Identify opportunities, barriers, and solutions for widespread implementation and impact, using the RE-AIM framework to address issues related to EMPathicO's Reach, Effectiveness, Adoption, Implementation, and Maintenance.

1.4 Definition of endpoints

A complete list of patient measures is provided in table 1 and practitioner measures in table 2 at the end of the document.

1.4.1 Definition of primary endpoint

For the MSK group, the two primary outcomes are **pain intensity and patient enablement**, each averaged over 6 months using a repeated measures approach, and we have allowed for two primaries in our sample size calculations. The outcomes will be reported separately.

For the non-MSK group, patient enablement will be the single primary outcome. We will also measure pain intensity in this group but will treat it as a secondary outcome.

Pain Intensity

Pain intensity will be measured using the 4-item pain intensity subscale from the Brief Pain Inventory (BPI)[1], on a numerical rating scale from 0 to 10. We will use *average pain in the last week* as the primary measure of pain intensity. Pain intensity will be assessed before the index consultation to provide a baseline, then again immediately (within 1-week), 1-month, 3-months, and 6-months post-consultation.

Patient Enablement

The Patient Enablement Index (PEI)[2] is a modified 7-point agree-disagree Likert response scale[3], including a Not Applicable option. Enablement will be assessed immediately (within 1-week), 1-month, 3-months, and 6-months post-consultation

1.4.2 Definition of secondary endpoints

Secondary outcomes capture patient satisfaction with the consultation (within 1 week immediately post-consultation) and health and quality of life changes (at 1-, 3- and 6-months post-consultation, see Patient Timelines Table).

Pain intensity

The 4-item aggregated pain intensity subscale (worst, least, average, now) will be used as a secondary measure of pain intensity.

Symptom Severity and Global Impression of Change

Two single item 7-point measures of Patient Global Impression of Symptom Severity and Patient Global Impression of Change[4] will be administered at baseline (symptom severity item only) and all post-consultation timepoints (both items).

Patient Satisfaction

The 21-item Medical Interview Satisfaction Scale (MISS)[5,6] measures patient satisfaction with the consultation, and includes three subscales (Professional Care, Depth of Relationship and Perceived Time).

Pain Interference

Pain interference will be measured with the 7-item pain interference scale from the BPI.

1.5 Analysis principles

All analyses will be reported according to CONSORT: extension to cluster randomised trials on planning, implementing and reporting statistical analyses[7].

2 Design considerations

2.1 Description of trial design

A cluster-randomized controlled two parallel groups superiority trial in primary care; general practices constitute the clusters, which will be randomized 1:1 to EMPathicO training versus usual care. All eligible practitioners (see below) within each cluster will be encouraged to undertake EMPathicO training (intervention) or will consult patients as usual (control); patients who present to participating practitioners will complete patient reported outcome measures at baseline and at four subsequent timepoints, assessing pain, enablement, and secondary outcomes.

Practices will be randomly allocated to one of two randomised arms:

1. Usual care
2. Access to EMPathicO digital e-learning package for practitioners

There are two groups of participants:

1. MSK group: consulting about MSK pain, with a score of 4 or more on Brief Pain Inventory at the index consultation
2. All-comers: consulting about something other than MSK pain, or consulting about MSK pain and with a score of less than 4 on the BPI at index consultation

2.2 Trial power and sample size

MSK Group

The minimum clinically important difference in the Brief Pain Inventory is around one point, with a standard deviation of 3.3. This is consistent with a standardised effect size of 0.3. For 90% power, two-sided alpha of 0.025 to account for the two primary outcomes, and a correlation between the 4 repeated measures (excluding baseline) of 0.7, a sample size of 214 per arm is required. We assume a conservative ICC of 0.03, at the upper 75% percentile of what has been observed in previous primary care trials. Assuming 20 patients per practice gives a design effect of 1.57. Allowing for 20% loss to follow up gives a total sample size of (214*2*1.57)/0.8=840 participants to be recruited from 42 practices.

All-comers Group

840 all-comer patients will give us 90% power (based on two sided alpha 0.05 and ICC as per the MSK group above) to detect a standardised effect size of 0.3 in the PEI, which is equivalent to a difference of 0.36 points (assuming SD=1.2, based on feasibility study).

Updated sample size calculation
Participants were recruited from 53 practices rather than 42 practices as originally planned, which reduced the average cluster size. Assuming 14 patients per practice gives a design effect of 1.39. Under the same assumptions as above, the total sample size is (214*2*1.39)/0.8=744 participants.

2.3 Randomisation details

An independent programmer produced a computer-generated allocation sequence with random block sizes of 4 and 6 and stratification by high/low deprivation (IMD score 5 or less/IMD score greater than 5) and large/small practice list size (greater than or equal to 7900/ less than 7900). The allocation sequence will be implemented using LifeGuide software. Patients and the trial statistician will be masked to intervention allocation.

2.4 Timing of planned analyses

Analyses will be carried out after 6 month follow up.

2.4.1 Interim analyses and early stopping

No interim analyses are planned

2.4.2 Stopping rules

No stopping rules planned

2.5 Final analysis

All primary and secondary analyses will take place when all patients have completed 6 months follow up, or have been lost to follow up after 6 months.

3 Statistical considerations

3.1 Definition of analysis populations

3.1.1 Intention-to-treat analysis population

Analysis will be by intention to treat (as randomised) regardless of any practice-level non-adherence to the intervention. All summaries and analysis will be on the ITT population unless otherwise specified.

3.2 Analysis software

Analysis will be carried out using Stata version 17 or higher.

3.3 Methods for handling data

3.3.1 Withdrawal from trial

All data up until the point of patient withdrawal from the trial will be used in analyses unless the patient withdrew consent and does not wish for the data already collected prior to withdrawal to be used for the trial.

3.3.2 Missing data

All available data will be analysed, with a sensitivity analysis using multiple imputation if appropriate. Linear mixed models and multiple imputation both assume the data are missing at random, therefore sensitivity analyses to data missing not at random will also be explored. This is set out in section 4.4 (primary analysis). The imputation model will include all variables in the analysis model and any variables which predict the missingness of the outcome.

3.3.3 Outliers

If outliers are found in regression modelling, then firstly the source data will be checked. If the source data shows that the data is correct, then the outliers may be excluded as a sensitivity analysis to explore any differences in results.

3.3.4 Assumption checking and alternative methods

Assumptions for linear mixed models (including normality, linearity and homoscedasticity) will be checked. If linear modelling assumptions are not met, another appropriate parametric distribution will be sought and a suitable generalised linear mixed model will be used.

3.3.5 Data transformations

Data transformations, such as log transformation, may be considered if data are highly skewed.

3.4 Definition of key derived variables

Patient enablement[2] is scored by taking the mean of 6 items on a scale 0 to 7, with higher values indicating greater enablement.

A secondary measure of pain intensity is the mean of the 4-item pain intensity subscale[1] (worst, least, average, now) with a score from ranging from 0 to 10 (higher scores indicating worse pain).

Patient global impression (PGI)[4] of symptom severity is scored from 1 (none) to 7 (extremely severe), and PGI of symptom change is scored from 1 (very much improved) to 7 (very much worse).

Pain interference is scored as a mean of seven items, each ranging from 0 to 10, with higher scores indicating greater interference.

Patient satisfaction (MISS)[5,6] is scored by calculating the mean of 21 items and reversing the scores of negatively phrased questions. The produces a score ranging from 1 to 7, with higher scores indicating greater satisfaction.

Perceptions of practitioner empathy (CARE)[8] is scored by adding all ten items, producing scores ranging from 10 to 50. Up to two 'Not Applicable' responses or missing values are allowable and are replaced with the average score for the remaining items. Otherwise, the score will not be calculated.

Perceptions of practitioner optimism is on a scale of 1 to 7 with higher scores indicating greater perceived optimism.

Treatment expectations (TEX-Q)[9] assesses expectations of treatment benefit, positive impact, adverse events, negative impact, process and behavioural control with a total of 15 items. An overall mean score of the TEX-Q can be calculated after reversing the harm expectation subscales (items 7–11), with higher values indicating more positive overall treatment expectations. In addition, the following subscales will be calculated: treatment benefit (1-3); positive impact (4-6); adverse events (7-9); negative impact (10-11); process (12-13); behaviour centre (14-15).

Anxiety and depression (HADS)[10] - Odd numbers: 1, 3, 5, 7, 9, 11 and 13 are Anxiety Questions; Even numbers: 2, 4, 6, 8, 10, 12 and 14 are Depression Questions. Scores for each subscale range from 0 to 21, with the following interpretation: 0-7 (Normal), 8-10 (Mild), 11-15 (Moderate), 16-21 (Severe)

Continuity of care (Patient doctor depth of relationship scale)[11] – a total score is calculated, which ranges from 0 (no relationship) to 32 (very strong relationship), provided 6 or more items are completed: Depth-of-relationship = (Mean score of completed questions/Maximum question range) × 32

Unless otherwise stated, one missing item will be allowed and replaced with the mean of the other items. If more than one item is missing, the score will not be calculated.

3.5 General principles for reporting and analysis

Analyses will, in general, be reported using a two-sided significance level of 5%, corresponding to 95% confidence intervals. The two primary outcomes in the MSK group will be reported using a significance level of 2.5%, corresponding to 97.5% confidence intervals. Results adjusting for stratification factors and baseline covariates will be reported. Descriptive statistics will be reported to 1 decimal place as number and percentage for categorical variables, and mean and standard deviation for continuous variables, or median and interquartile range for variables with a skewed distribution. The randomised arms will be labelled Usual Care and Intervention.

4 Planned analyses and reporting

4.1 Disposition of the study population

CONSORT flow diagram (following CONSORT guidelines) which should include flow of practices and individual participants through each stage. For each randomised arm report the numbers of practices and participants randomly assigned, receiving intended treatment, completing follow-up, and analysed for the primary outcome. Include the number of clusters, average cluster size, and range of cluster size at each stage. Numbers of participants will be reported separately by the MSK and all-comer groups. Describe protocol deviations from study as planned, together with reasons.

4.2 Protocol deviations

A protocol deviation is an unanticipated or unintentional divergence or departure from the expected conduct of a study inconsistent with the protocol, consent documents or other study procedures. Of particular importance are major deviations (violations) which may expose participants to increased risk; compromise the integrity of the entire study or affect participant eligibility. Protocol deviations will be listed with information on treatment arm and the type of deviation. Full details of the protocol deviations will also be listed. Failure to engage with the randomised intervention and missing follow ups will not be taken as protocol deviations.

4.3 Baseline and demographic characteristics

Baseline patient measures will be tabulated using appropriate descriptive statistics by randomised arm, separately for the MSK group and all-comer participants. These consist of patient socio-demographic measures (age, gender, index of multiple deprivation, ethnicity, education level, work status) and baseline patient outcome measures listed in Table 1.

Baseline practitioner demographic (age, gender, ethnicity, years qualified, profession) and outcome measures (listed in Table 2) will be summarised descriptively by randomised arm.

Practice-level measures (practice list size, practice index of multiple deprivation) will also be summarised by randomised arm.

4.4 Primary endpoints

MSK group

For the MSK group, the primary analyses for the pain (BPI ‘average’ pain) and enablement (PEI) scores will be performed using a generalized linear mixed model (GLMM) framework with observations at 7 days, 1-, 3-, and 6-months (level 1) nested in participants (level 2) and

participants nested in practices (level 3). This model will be used to compare the mean pain and enablement scores across follow-up time points between participants receiving the TIP intervention and those receiving usual care. Results will be reported adjusting for stratification factors (high/low practice deprivation, large/small practice size) and baseline BPI (average pain) score. As there may not be a constant treatment effect over time, a treatment/time interaction will also be included. An unstructured covariance matrix will be used to model covariance between repeated measures on each individual. The treatment effect estimates will be reported with 97.5% confidence intervals, as well as the intra-cluster correlation for each primary outcome. The main effect will be reported, unless there is a significant treatment/time interaction, in which case treatment effects at each timepoint will be reported. Clinical effectiveness will be concluded if the point estimate of the treatment effect favours the intervention and the 97.5% CI excludes zero, for either pain or enablement outcomes.

The primary analysis uses all available data in the mixed model and implicitly assumes that missing outcome scores are missing at random (MAR) given the observed data. If appropriate, a sensitivity analysis using multiple imputation with chained equations will be performed. The imputation model will include all variables in the analysis model and any baseline covariates predictive of the missingness of the outcomes. Imputation will be performed separately by randomised arm and then combined. Sensitivity analyses to the MAR assumption will include a pattern mixture model approach: assuming missing pain scores are on average 1 point better or worse than the observed scores, and missing enablement scores are on average 1 point better or worse than the observed scores.

All comers group

For the all-comers group, the same model will be fitted for enablement as above. Clinical effectiveness will be concluded if the point estimate of the treatment effect favours the intervention, and the 95% confidence interval excludes zero.

4.5 Secondary endpoints

The following secondary and process outcomes will be analysed for MSK and all-comers groups separately (pain intensity and pain interference will only be available for a subset of all-comers)

For secondary outcomes, 4-item pain intensity, PGI symptom severity and PGI symptom change (measured at 7 days, 1-, 3-, and 6-months), the analyses will use a similar linear mixed modelling approach to the primary analysis and control for baseline score and the same baseline covariates.

Pain interference at 1 and 6 months will be analysed using linear mixed model, adjusting for baseline score and the same baseline covariates as in the primary analysis. Patient satisfaction at 7 days will be analysed using linear mixed model.

4.6 Process measures

4.6.1 Patient measures

Perceptions of practitioner empathy (CARE), perceptions of practitioner optimism, treatment expectations (TEX-Q), and continuity of care (PDDOR), at 7 days, will be analysed using linear mixed model, adjusting for the same variables as in the primary analysis. Anxiety score and depression score (HADS) at 7 days will be analysed using linear mixed model, adjusting for the same covariates as in the primary analysis. Pain medication will be summarised descriptively by randomised arm.

4.6.2 Practitioner measures

Practitioner self-efficacy for conveying empathy and practitioner self-efficacy for conveying optimism at 8 and 34 weeks will be analysed using linear regression, adjusting for stratification factors and baseline self-efficacy.

Practitioner outcome expectancy and intentions to set goals will be summarised descriptively for the intervention arm only.

4.7 Additional analyses

4.7.1 Subgroup analyses

Exploratory subgroup analyses will be performed by repeating the primary analyses and including a treatment by covariate interaction term for the following subgroups:

- Reasons for consulting:
    - Investigations, test results, medication review;
    - Mental health problems;
    - Any ICD-10 subgroup accounting for at least 10% of sample
- Multimorbidity – 2 or more comorbidities / 1 or fewer comorbidities
- Pain score at baseline – mild (1 to 3)* / moderate (4 to 7) / severe (8 to 10)
- Index consultation modality – in person / telephone
- Type of practitioner – GP / nurse/physio/other (subgroups may need to be combined)

- Age – 18 to 45 / 45 to 65 / above 65
- Gender – male / female
- Index of multiple deprivation decile – higher (1 to 5) / lower (6 to 10)

*For the pain outcome, subgroup analyses may only be carried out among the MSK group (baseline pain ≥4) and those providing pain scores in the all-comers group (baseline pain < 4). For the enablement outcome, data from the MSK and all-comers groups can be combined.

4.7.2 Further analyses

The health economic analysis plan and the process evaluation plan will be outlined in separate documents.

4.8 Safety reporting

An adverse event (AE) is any untoward medical occurrence in a trial participant which does not necessarily have a causal relationship with this treatment. An AE can therefore be any unfavourable and unintended sign (including abnormal laboratory finding), symptom or disease. A serious adverse event (SAE) is any AE that results in death, is life-threatening, requires hospitalisation or prolongation of existing hospitalisation, results in persistent or significant disability or incapacity, consists of a congenital anomaly or birth defect, or other medically important condition. SAEs will be summarised descriptively by randomised arm.

5 Tables, listings and figures

Table 1. Patient-Reported Characteristics, Outcomes and Process Variables

Variable	Measure	Items	Measurement Timings				
			<-7d	<7 d	+1 m	+3 m	+6 m
Primary Outcomes							
Pain intensity (pain sample)	Pain intensity subscale from the BPI	4	x	x	x	x	x
Patient enablement	Modified PEI[77]	6		x	x	x	x
Secondary Outcomes							
Patient global impression of symptom severity	Single item[78]	1	x	x	x	x	x
Patient global impression of symptom change	Single item[78]	1		x	x	x	x
Pain interference	Pain interference subscale from the BPI[76]	7			x		X
Patient satisfaction	MISS for UK general practice[79]	21		x			
Adverse events	Bespoke self-report item	1			x	x	x
Health Economics							
Health-related quality of life	EQ-5D-5L and EQ-VAS[80]	6	x		x		x
Capability wellbeing	ICECAP-A[81 82]	5	x		x		x
Healthcare utilization	ModRUM core module[83]	12		x		x	x
Prescribed medications	ModRUM depth questions[83]	1				x	x
Personal expenses	Bespoke self-report item	3				x	x
Productivity	WPAI:GH	6				x	x
Process Measures							
Perceptions of practitioner empathy	CARE[38]	10		X			
Perceptions of practitioner optimism	Bespoke item	1		X			
Treatment expectations	Treatment expectation questionnaire TEX-Q[84]	15		X			
Anxiety	HADS[85 86]	7		X			
Continuity of care	Patient-Doctor Depth of Relationship Scale[87]	9		X			
Depression	HADS[85 86]	7		X			
Sociodemographic Characteristics							
Age, gender, ethnicity		3	x				
Index of Multiple Deprivation	Postcode	1	x				
Health Characteristics							
Reasons for consulting		1		x			
Comorbidities		1		x			
Index consultation modality		1		x			

Table 2. Practitioner-Reported Characteristics, Outcomes and Process Variables

Practitioners	Variable	Measure	Items	Measurement Timings			
				Baseline	+2wk	+8wk	+34wk
All	Characteristics (age, gender, ethnicity, years qualified, profession)	Bespoke	5	x			
All	Practitioner self-efficacy for conveying clinical empathy	Bespoke, from feasibility study	7	X		X	x
All	Practitioner self-efficacy for conveying realistic optimism	Bespoke, from feasibility study	5	x		X	x
Intervention arm only	Practitioner outcome expectancy for implementing goals set during EMPathicO training	Bespoke, from feasibility study	16	X		X	x
Intervention arm only	Practitioner intentions to implement goals set during EMPathicO training	Bespoke, from feasibility study	3	X		X	x
Intervention arm only	Practitioner intervention usage	LifeGuide data	N/A			X	X
All	Practitioner-reported other training	Bespoke	1			x	x

6 References

1. Keller S, Bann CM, Dodd SL, et al. Validity of the Brief Pain Inventory for use in documenting the outcomes of patients with noncancer pain. Clin J Pain 2004;20(5):309-18.
2. Howie JG, Heaney DJ, Maxwell M, et al. A comparison of a Patient Enablement Instrument (PEI) against two established satisfaction scales as an outcome measure of primary care consultations. Fam Pract 1998;15(2):165-71.
3. Preston CC, Colman AM. Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent preferences. Acta Psychol (Amst) 2000;104(1):1-15. doi: https://doi.org/10.1016/S0001-6918(99)00050-5
4. Fischer D, Stewart AL, Bloch DA, et al. Capturing the Patient's View of Change as a Clinical Outcome Measure. JAMA 1999;282(12):1157-62. doi: 10.1001/jama.282.12.1157
5. Wolf MH, Putnam SM, James SA, et al. The Medical Interview Satisfaction Scale: development of a scale to measure patient perceptions of physician behavior. J Behav Med 1978;1(4):391-401. [published Online First: 1978/12/01]
6. Meakin R, Weinman J. The 'Medical Interview Satisfaction Scale' (MISS-21) adapted for British general practice. Fam Pract 2002;19(3):257-63. doi: 10.1093/fampra/19.3.257
7. Campbell MK, Elbourne DR, Altman DG; CONSORT group. CONSORT statement: extension to cluster randomised trials. BMJ. 2004 Mar 20;328(7441):702-8. doi: 10.1136/bmj.328.7441.702. PMID: 15031246; PMCID: PMC381234.
8. Mercer SW, Maxwell M, Heaney D, et al. The development and preliminary validation of the Consultation and Relational Empathy (CARE) measure: an empathy-based consultation process measure. Fam Pract 2004;21 699-705.
9. Alberts J, Löwe B, Glahn MA, et al. Development of the generic, multidimensional Treatment Expectation Questionnaire (TEX-Q) through systematic literature review, expert surveys and qualitative interviews. BMJ Open 2020;10(8):e036169. doi: 10.1136/bmjopen-2019-036169
10. Bjelland I, Dahl AA, Haug TT, et al. The validity of the Hospital Anxiety and Depression Scale: An updated literature review. J Psychosom Res 2002;52(2):69-77.
11. Ridd MJ, Lewis G, Peters TJ, et al. Patient-Doctor Depth-of-Relationship Scale: Development and Validation. The Annals of Family Medicine 2011;9(6):538. doi: 10.1370/afm.1322

7 SAP revision history

Version number	Revision history	Author	Date