Healthy London Partnership

DIALOG scale – analytical framework for mental health services

Rahul Bhattacharya, Stefan Priebe and Victoria Bird, January 2019

What does DIALOG measure?

DIALOG has 11 items. Each item is rated on a Likert scale with 1 (=totally dissatisfied), 2 (=very dissatisfied), 3 (=fairly dissatisfied), 4 (=in the middle), 5 (=fairly satisfied), 6 (=very satisfied) and 7 (=totally satisfied). NB for data capture and reporting purposes 0 = clinician left blank, whilst 8 = patient left blank / refused to answer.

The first eight items ask to rate the satisfaction with different domains of life. Their mean score reflects the overall score for subjective quality of life as a Patient Rated Outcome Measure (PROM).
The final three items ask to rate different aspects of treatment. Their mean score reflects the overall score for treatment satisfaction as a Patient Reported Experience Measure (PREM).

DIALOG also allows the interpretation of scores on every single item of interest. As far as possible, results should always be presented for all 11 items. For this one can use a graph, e.g. showing the mean scores with horizontal bars as they are displayed in the DIALOG app.

What is a positive or negative score?

All scores – including all mean scores for groups of patients – refer to the underlying 1 to 7 rating scale and can therefore be directly interpreted. Thus, every score below 4 – as a mean score or on single items - reflects explicit dissatisfaction, 4 is the neutral middle, and every score above reflects 4 explicit satisfaction. The degree of satisfaction or dissatisfaction varies in line with the labels of the rating scale of the items (see above).

When should ratings be obtained?

For each patient and treatment episode, there are entry and exit ratings, and for longer term treatments there are review ratings in between. In principle, ratings should be obtained at the beginning of any new treatment episode (e.g. admission to in-patient treatment, home treatment service (HTT) or other community service; referral to new provider organisation) and at the end of that treatment episode (e.g. discharge from that service or provider organisation). For acute treatment the expectation is that initial ratings are obtained at the latest 48 hours after admission, this includes admission to the ward or to the HTT. Research shows that patients are able to provide meaningful ratings including on their treatment satisfaction already after one day in a new treatment (e.g. one day after hospital admission). High symptom levels may make any structured conversation impossible, but – when patients are willing and able to do the ratings – symptom levels normally do not invalidate the ratings. High levels of depression are likely to lead to more negative ratings, as low mood and negative appraisals of all aspects of life are generally linked. Yet, patients with marked depressive symptoms still distinguish in their ratings between different domains of life and different treatment aspects. For patients treated

within the community, the initial rating is expected to be obtained within the first or at the very latest, second meeting.

For treatments over longer periods of time including within the community, review ratings are required, ideally at fixed time intervals. The fixed time intervals for acute treatment (in-patient and crisis teams) can be shorter (e.g. every four weeks), for on-going treatment in community services and out-patient clinics they should be not longer than every six months .

Evaluating treatment

When evaluating treatment, for quality of life, change scores are considered, i.e. whether quality of life has improved during treatment or not. For this, either mean scores across the eight rated life domains or ratings of single life domains can be selected, depending on the focus and context of care, the situation of the patients and the purpose of the evaluation. For example, some evaluations may focus on domains with which patients express particularly low satisfaction at intake, whilst other evaluations may put emphasis on single domains that the given service model targets.
Mean scores can be calculated even when one item is missing (that item is then ignored when calculating the mean). Yet, when more than one item is missing, mean scores may be substantially affected.
Changes are the differences between ratings at two time points (e.g. admission and discharge) and can be displayed graphically. High initial levels of satisfaction scores provide less scope for improvement (so called ceiling effect) whilst very low scores make improvements more likely (so called regression to the mean).

For treatment satisfaction, absolute scores at a given time point are more relevant than changes over time, i.e. whether patients are or are not satisfied with treatment irrespective of how their satisfaction was in the past.

Interpretation of scores for individual patients

All scores of individual patients can be used for an assessment of the personal problems and areas of strengths. For subjective quality of life, scores below 4 - reflecting explicit dissatisfaction - should receive particular attention.
Most mean scores are likely to be above 4, and mean scores above 5 (i.e. more than fairly satisfied) should raise the question why precisely the patient is receiving secondary care and what should be achieved for the patient in further treatment.

When evaluating treatment, any improvement in subjective quality of life (including one point on only one item) should be seen as an achievement as it reflects a meaningful increase of satisfaction in at least one life domain. However: for patients in long-term care, it may be unrealistic to expect consistent and on-going improvements, in which case scores have to be interpreted in consideration of the personal context of the patient.

For treatment satisfaction, all ratings should consistently stay above 4 demonstrating an at least fair degree of satisfaction with the three main treatment aspects (which is a realistic expectation given that treatment satisfaction in mental health care is usually high).

Interpretation of scores for all patients in a service

Scores of single items of patients in a service can be shown as means (=average satisfaction) and/or as percentages of patients who score 4 and lower (=% of patients explicitly dissatisfied with that domain). The item scores can be used for an assessment of the most prevalent and/or most severe problems of patients and of areas of strengths of the patients in the service. Comparisons with other services can help to identify specific tasks in the given service that are different from the general challenges to all services.

When considering changes over time in the evaluation of services, one can consider the global mean score as well as single items and this will depend on the precise purpose of the evaluation. If one evaluates full services with variable treatment tasks for a range of patients, it is reasonable to explore first the global scores before considering the ratings of single life domains. This applies to most services in most situations. However, there may also be specific questions, e.g. on how services affect the wide spread dissatisfaction of patients with their physical health. Analysing the global score provides a general evaluation, whilst the analysis of single items provides more specific information about what exactly services do and do not achieve.

For interpreting changes over time, the question is which change is meaningful. As a principle, every single improvement in each area is worth having. Changes can be shown as changes in mean scores or changes in the percentage of explicitly dissatisfied patients. The second approach shows how many patients have improved overall and in each domain, and both means and percentage of dissatisfied patients can be shown in graphs).

The average subjective quality of life score of all patients in a service should not improve, as patients with better quality of life will be discharged and new patients with lower quality of life be admitted. Yet, for the same patients scores should improve. This certainly applies to acute and short-term treatments, whilst the options for ongoing improvement in long-term treatments are limited. A change of overall mean scores of >0.125 reflects an average improvement of at least one scale point in at least one life domain and may be seen as a guide for an overall meaningful improvement.

Treatment satisfaction scores should consistently stay above 4 and the percentage of patients with explicit dissatisfaction with any treatment aspect should be kept as a minimum. Still, comparisons of the level of treatment satisfaction across comparable services may help to identify good practice and potentials for improvements.

Interpretation of scores for all patients in an organisation

The interpretation of scores across all patients in an organisation follows the same principles as the interpretation for all patients in a team. Since the patient groups are very large, differences between mean scores tend to become smaller (=tendency towards the mean) and percentages of explicitly dissatisfied or very satisfied (particularly for aspects of treatment satisfaction) may be more informative than mean scores.