**Online supplemental data for the MONARCA II trial**

***e1***

**METHODS**

*Intervention*

Regardless of randomization allocation (intervention group or control group), all patients were offered to borrow an Android smartphone free of charge during the nine months trial period. The patients, regardless of randomization allocation, who were using a smartphone not capable of collecting automatically generated objective data (iPhone) could choose to continue to use their own smartphone.

The intervention group

*The self-monitoring items:* The following self-monitored measures were evaluated on daily basis: Mood (scored from depressive to manic on a scale from -3, -2, -1, -0.5, 0, +0.5, +1, +2, +3), sleep duration (number of hours slept per night, measured in half-hour intervals), medicine intake (taken as prescribed/ taken with changes (if changes, the patients are asked to specify these)/not taken), activity level (scored from very low to very high on a scale from -3, -2, -1, 0, +1, +2, +3), mixed mood (yes/no), irritability (scored from not present, present to some degree or present on a scale from 0, 1, 2), anxiety (scored from not present, present to some degree or present on a scale from 0, 1, 2), cognitive problems (scored from not present, present to some degree or present on a scale from 0, 1, 2), alcohol consumption (number of units consumed per day, 0 to +10 scale), stress (scored from not present, present to some degree or present on a scale from 0, 1, 2), menstruation for women (yes/ no), individual early warning signs (yes/ no), a number (unlimited) of personal parameters (created by the patients themselves), and a free-text note. After midnight, the entered self-monitored measures of illness activity were “locked” and further changes could not be made. If the patients forgot to evaluate it was possible to enter and evaluate retrospectively for up to two days.

*The objective smartphone data:* The following objective smartphone data were included in the clinical feedback loop: Phone usage measured as the amount of time the smartphones screen were turned on/off; Social activity measured as the number of in- and outgoing phone calls and text messages, the duration of in- and outgoing phone calls, the length of the text messages, and the time of the day when the phone calls and/or text messages were made/send/ or received; Physical activity measured by the step counter in the smartphones; Mobility based on the location estimation available in the smartphones.

*The integrated clinical feedback loop:* The clinical feedback loop between patients and clinicians comprised a study nurse who examined the collected data on a web-page two to three times a week, or more often on patients where it is deemed necessary. At the very first meeting between the study nurse and the patient an alliance and a concordance on the patients’ goal during the trial were established. In collaboration and based on the patients’ needs, the study nurse and the patient made crisis plans on actions to take in case of depression or (hypo)mania and plans for preventing future depression and (hypo)mania continuously during the trial. Further, activity planning, weekly plans and coping strategies were made and revised continuously. Based on the patients’ current state, the study nurse continuously provided psychoeducation, supportive therapy and advise on the use of *pro necessitate* medication.

1) The clinical feedback loop on patient-reported measures: Regardless the choice of smartphone a feedback loop on the subjective measures was established. A standard of scoring thresholds for when the study nurse initially should react was made. For example, the study nurse reacted if the patients registered ≥-2 on the mood item for two days or more, or if the patients registered changes in their sleep patterns of 1 hour or more for more than three days. Lack of self-monitoring data provided by the patient was interpreted a sign of deterioration. Following a run-in phase of approximately two to four weeks of self-monitoring, the patients and the study nurse individualized the thresholds for when reaction should be made. The study nurse and the patients agreed on a concordance status in a) the patients most important items for identifying prodromal symptoms of depression as well as (hypo)mania b) the threshold for future early warning signs c) actions to be taken in case of depression or (hypo)mania.

2) The clinical feedback loop on patient-reported measures and objective smartphone data: A feedback loop integrating patient-reported measures and objective smartphone data on measures of illness activity was established for patients using smartphones capable of collecting objective smartphone data on measures of illness activity. The feedback loop integrates both patient-reported measures and objective smartphone data on measures of illness activity in a flexible and adjustable model (a learning system) resulting in prediction analyses of the collected data providing messages for both the patients and the study nurse such as: “you should contact the study nurse”.

Actions by the study nurse as part of the clinical feedback loop in the intervention group: In the case of signs of deterioration of a patient the study nurse:

a) Contacted the patient and gave advice on how to handle the situation either over the phone or by consultation in the clinic. This included clinical evaluations of the patients’ current state and evaluation of suicidal risk.

b) If the first action was not enough – the study nurse asked the patient to contact his/her usual physician or other clinician.

c) If the above actions were not enough, or if contact to the patient was not possible – the study nurse contacted the patient’s usual physician or other clinician.

d) If acute deterioration and/ or severe symptoms – the study nurse contacted the psychiatric emergency service in Copenhagen.

***e2***

**RESULTS**

Overall, a total of 54 patients included in the MONARCA II trial borrowed a smartphone (Samsung Galaxy or LG) during the trial period (intervention group, n=33; control group, n=21) (the main reason being: had an older version of smartphone and wanted to use a new one).

The intervention group

During the nine months study period patients in the intervention group adhered to the daily self-monitoring 72.6% (196 days) of the days. The patients expressed that the self-monitoring system was supportive, useful, quick and easy to use with a low level of intrusiveness. None of the patients included in the intervention group expressed that they felt watched, and none of the included patients were uncomfortable with having the objective smartphone data collected but saw it as a safety net.

During the trial, the study nurse conducting the clinical feedback loop had a mean number of times in contact with the patients of 14.6 (range 2 to 47). The main communication form between the patient and the study nurse was via phone calls and text messages. The main part of the meetings between the patients and the study nurse at the clinic were scheduled in advance (mean number of visits per patients during the trial 2.6 (0 to 20). Very few patients communicated by e-mail (mean number of e-mails per patient: 0.6; range 0 to 4). During the trial a total of 66 notifications based on automatic prediction analyses on the collected data were sent to the study nurse and the patient. Of these, a total of 31 provided information to the study nurse regarding current depressive symptoms, and thus provided opportunities for early intervention. In 35 of the notifications, the patient was stable and did not have signs of deterioration.

Due to unforeseen technical issues with the real-time mood forecasting based on patient-reported data as well as objective smartphone data, the forecasting part of the clinical feedback loop were not fully functioning at all times during the trial (approximately 1/3 of the time). Thus, at times the patients in the two intervention groups (Android or iPhone) received quite similar intervention elements during periods of the trial, and the main part of the clinical feedback loop was based on the patient-reported data and clinical evaluation of these.

We did not have information regarding the economic income, but there were no statistically significant differences in employment status or education between any of the three groups (p-value>0.82). There was no statistically significant difference in the retention rate between the three groups (p-value>0.54).