BMJ Open | Lessons learnt from a cluster-randomised trial evaluating the effectiveness of Self-Management Support (SMS) delivered by practice nurses in routine diabetes care

Anneke van Dijk-de Vries,[1,2] Marloes A van Bokhoven,[1] Bjorn Winkens,[3] Berend Terluin,[4] J André Knottnerus,[1] Trudy van der Weijden,[1] Jacques Th M van Eijk[2]

For numbered affiliations see end of article.

Correspondence to
Anneke van Dijk-de Vries;
anneke.vandijk@
maastrichtuniversity.nl

ABSTRACT

Objective: To evaluate the effectiveness of biopsychosocial Self-Management Support (SMS) delivered by practice nurses in routine diabetes care.

Design: A pragmatic cluster-randomised controlled trial within a hybrid effectiveness-implementation study design. Practice nurses were cluster-randomised.

Setting: A regional care group in the Netherlands consisting of 77 family practices. The study involved practice nurses (n=40) providing care to approximately 4000 patients with diabetes.

Participants: Patients with type 2 diabetes (n=264) selected by a self-administered questionnaire aimed at measuring emotional distress and diabetes-related reduced daily functioning.

Intervention: Practice nurses in the intervention arm (n=19) were trained to integrate SMS into their routine consultations. SMS included detection of patients with emotional distress and reduced daily functioning, and supporting them when needed through problem solving and reattribution techniques. Practice nurses in the control arm (n=21) provided usual care.

Main outcome measures: The primary outcome measure was a dichotomised score on a Visual Analogue Scale that measured the perceived effect of diabetes on daily functioning. Secondary measures included patients' diabetes-related distress, quality of life, autonomy and participation, self-efficacy, self-management and glycaemic control. Outcomes were measured at baseline and at 4-month and 12-month follow-ups.

Results: Only 16 of the 117 patients in the intervention arm (14%) who were found eligible by the posted research-driven screening questionnaire were detected by their practice nurses. Extra consultations for the self-management support were delivered to only 11 study participants. In the control arm, 147 patients received usual care. Multilevel analyses showed no significant differences in outcomes between the intervention and control arms.

Conclusions: SMS in its present form was not effective. The research-driven screening to select trial participants appeared to be inconsistent with nurse-led

Strengths and limitations of this study

- A thorough cluster-randomised trial to evaluate the effectiveness of an evidence-based intervention integrated into routine primary care.
- Limited interference of the researcher in the routine care setting.
- Low exposure of study participants to the complete intervention.
- Absence of a pilot of the detection method prior to implementation.

detection in routine practice. Adequate follow-up moments need to be built in to overcome barriers resulting from tension between the implementation and effectiveness parts of hybrid studies.

Trial registration number: Current Controlled Trials NTR2764.

INTRODUCTION

Patients with chronic diseases such as type 2 diabetes mellitus have to deal with challenging day-to-day management tasks regarding the medical, emotional and social consequences of their chronic condition.[1 2] Research has shown a reciprocal relationship between the emotional health status of people with diabetes and their medical self-management: emotional distress may interfere with control over the disease, whereas poor control over the disease can lead to emotional distress.[3–6] Studies highlight the need to support the medical as well as the emotional and role management tasks in newly diagnosed patients and patients who face the longer-term consequences of their chronic condition.[7–9]

In European countries, most patients with diabetes receive follow-up care in the primary care setting by nurses.[10] Practice nurses (PNs) in the Netherlands work according to guidelines that focus on medical and behavioural management.[11] [12] Furthermore, the financial reimbursement of care and patients' health outcomes are determined by biomedical targets.[13] Consequently, the psychosocial aspects of diabetes care are not systematically incorporated in clinical practice.[10] [13–15]

In collaboration with a regional care group of general practitioners (GPs) in the Netherlands and a health insurer, the SMS ('Self-Management Support') implementation project was initiated to realise a shift from biomedically oriented care towards a biopsychosocial approach in diabetes care.[16] The starting point was a nurse-led minimal psychological intervention. A previous randomised trial has shown that this intervention was cost-effective for patients with diabetes with minor to moderate depression: 9 months after receiving the intervention, depressive symptoms were significantly lower, there was a positive effect on patients' quality of life, and patients experienced less anxiety, possessed more self-efficacy skills, demonstrated better glycaemic control and showed more participation in comparison to control patients.[17–19] There was an implementation momentum. The health insurer promised to pay the costs for the extra care, and the care group gave a commitment to take care of training facilities and integration of SMS parameters into Electronic Medical Records. Some adjustments to the nurse-led intervention were crucial. Where the original intervention had been delivered at patients' homes by specifically trained nurses from the research team who selected eligible patients by means of an elaborate diagnostic procedure, in SMS it was provided by PNs as a structural part of their consultations delivered in the family practice. For the identification of eligible patients, a simple detection method became an inherent part of SMS. The eligibility criteria changed from having a mild to moderate depression towards suffering from both emotional distress and interference in their daily functioning due to the burden of diabetes. This focus on daily functioning instead of a diagnosis of depression was supposed to fit the primary care setting. Patients who were detected by the PN received the nurse-led minimal psychological intervention or were referred to the GP to see whether more specialised care would be required, depending on the severity of the symptoms.

SMS was evaluated by means of a type 2 hybrid effectiveness-implementation study design, in which the regional implementation strategies and the effectiveness of SMS were evaluated simultaneously.[20] This paper will focus on the effectiveness part of the SMS project, though it is closely connected to the implementation part. This paper reports the effectiveness of SMS integrated by PNs into routine diabetes consultations regarding patients' daily functioning, emotional health, quality of life, autonomy and participation, self-efficacy and self-management skills, and blood glucose levels.

METHODS
Study design
A two-arm, pragmatic, cluster-randomised controlled trial was conducted with PNs as the unit of randomisation and their patients with type 2 diabetes mellitus as the unit of analysis. Measurements were at baseline, at 4 months and at 12 months. The study was conducted between November 2011 and February 2013 and was registered in a Dutch public trial registry [NTR2764]. The study protocol has been published elsewhere.[16]

Study participants
All the family practices of one regional organisation of GPs (a so-called 'care group') in the South of the Netherlands were eligible to participate in the study. GPs were asked to participate in the SMS project. Their PNs for diabetes care were randomly assigned to an intervention or control arm. PNs in the intervention arm were trained in SMS, whereas PNs in the control arm provided usual diabetes care, conforming to the Dutch guidelines. Owing to the variety of PNs working solo versus in a team, and in one versus more practices, PNs were stratified into working alone in a practice, working in a team and working in different settings. The randomisation was performed by an independent research assistant who used a random number seed computer program to assign PNs to study arms, assuming an allocation ratio of 1:1. PNs working together in a practice were clustered for being randomised to the same trial arm to avoid the risk of contamination if the SMS and usual care would be delivered in the same family practice. Regarding the analysis, a team of PNs was seen as one cluster if patients could receive diabetes care from these PNs alternately.

In the selection of patients for the effectiveness trial, the following issues were crucial. First, we wanted to select patients who would receive the complete intervention including detection and follow-up. Furthermore, patients from the intervention and control arms should be selected in a similar way, without interfering in their routine care. We set up a screening procedure by means of written questionnaires sent to patients' home addresses to identify patients with actual problems of daily functioning and emotional distress. Patients with a clinically established diagnosis of type 2 diabetes mellitus were sent a letter by their GPs to introduce the SMS project. The enclosed self-administered questionnaire consisted of the screening instruments that intervention PNs would also apply for SMS in their routine practice. It included the 'Daily Functioning Thermometer' (DFT), which is a Visual Analogue Scale to measure how suffering under the burden of diabetes affects patients' perceived functioning in everyday life. Patients were asked to indicate a position between 0 (no burden

at all) and 10 (extreme burden) at a continuous vertical 10 cm line. A score of 4 was chosen to differentiate between patients who could benefit from support in their self-management skills (score >4) and patients who had found a satisfactory way to live with the consequences of their diabetes (score ≤4). The DFT has been developed for the purpose of this study. It is comparable to the Distress Thermometer, a validated questionnaire in the care for patients with cancer,[21] and a validated Visual Analogue Scale regarding the worst (0) and the best possible life (10) for adolescents with type 1 diabetes mellitus.[22] Next to the DFT, patients were asked to complete the three-item Distress Screener (DS), which is a quick-scan instrument for emotional distress and an indicator of potential underlying severe mental health problems.[23] Patients with score DFT >4 and DS >3 were eligible for trial participation. The time interval between the researcher-driven self-administered screening questionnaire and the face-to-face nurse-led detection procedure in the consultation room needed to be as short as possible. However, intervention patients should complete the baseline measurement before SMS is applied. Owing to the logistics, patients received the posted screening questionnaire 4–6 weeks before a planned diabetes consultation. Reminders were sent at least 3 weeks before a consultation. Patients were asked to return the completed questionnaire to the research centre and to give informed consent to be approached if they were found to be eligible for research follow-up measurements. Within a week after receiving the screening questionnaire at our research centre, eligible patients were sent an invitation letter together with an informed consent form for trial participation and the baseline measurement. Patients who gave informed consent knew whether they would receive an addition to their usual care or not. No details were given about the content of the intervention.

Intervention: SMS in routine care

During three 8 h training sessions, PNs in the intervention arm were trained to integrate the detection and follow-up phase of SMS into their daily practice. The training sessions were followed up by booster sessions to maintain and improve PNs' skills in SMS. PNs attended these booster sessions three or four times during the year of follow-up.

After the training sessions, PNs started to integrate SMS into their routine care practice. PNs were blinded regarding the outcomes of the recruitment procedure and study participation of their patients. They applied SMS in all their consultations with patients with diabetes. Financial reimbursement for the extra time spent on SMS was added to the bundled payment arrangement for diabetes care.[24]

SMS included a detection and follow-up phase. The flow chart of SMS is presented in figure 1. The detection phase of SMS started by exploring whether a patient experienced problems in daily life. PNs then applied the DFT by showing patients a vertical line ranging from 0 to

10 with 0 indicating no burden at all and 10 indicating extreme burden, and asking them to rate the burden of diabetes in their daily life. PNs also verbally administered the three questions of the DS.[23] Patients with score DS >3, which indicates high risk for underlying mental health problems, were asked to complete the Four-Dimensional Symptom Questionnaire (4DSQ)[25] immediately after the consultation. This self-report instrument is widely used in Dutch primary care to distinguish non-specific distress from depression, anxiety and somatisation. It differentiates between mild, moderate or severe symptoms. The completed 4DSQ was returned to the PN, who computed the sum scores on each domain.

The follow-up phase of SMS was based on the outcomes of the DFT and 4DSQ. Patients registering a 'severe' score on at least one of the four subscales of the 4DSQ were referred to the GP for further diagnostics and treatment.

Patients who experienced problems of daily functioning (DFT >4) and emotional health problems (moderate scores on at least one subscale of the 4DSQ) were offered consultations for SMS. These extra consultations delivered by PNs were aimed at supporting patients in their day-to-day management of diabetes and its emotional and social consequences. The intervention strategy derived from the principles of learning theory has been described elsewhere.[26] PNs supported patients in the processes of defining problems and finding solutions themselves, by applying problem-solving and reattribution techniques. Problem-solving consists of seven stages that efficiently address problems and their possible solutions.[27] The reattribution technique was applied to challenge patients to link feelings and cognition to consequent behaviour. Patients could use information from a diary in which they recorded symptoms, thoughts, worries, feelings, and behaviour. Both problem solving and reattribution techniques were intended to result in action plans indicating how patients would achieve their personal goals.[28]

Measurements effectiveness trial

We used postal questionnaires for patient measurements. Patient characteristics regarding their age, sex, year of diagnosis, treatment of diabetes (insulin therapy or tablets), marital status, work status and educational level were assessed at baseline.

The glycaemic control of patients was measured during diabetes consultations. These measurements were extracted from the electronic patient databases.

In the year following the training, PNs in the intervention arm recorded process measures and outcomes of SMS on a specific registration form. These data were used to evaluate whether patients were exposed to SMS.

Outcomes

The effect of SMS on patients' daily functioning was measured by means of the DFT. The DFT was dichotomised to compare the number of patients who were improved in their daily functioning and had apparently found a

Figure 1 Self-Management Support (SMS) as implemented in routine care. DFT, Daily Functioning Thermometer; DS, Distress Screener; 4DSQ, Four-Dimensional Symptom Questionnaire; GP, general practitioner.
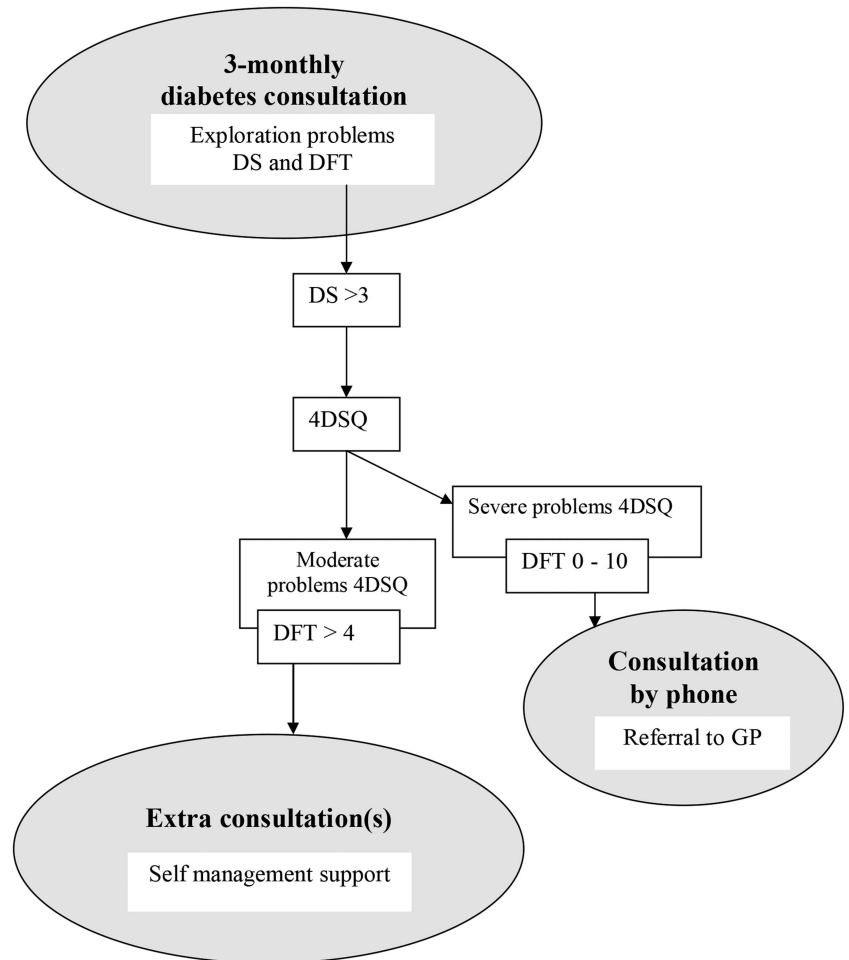
satisfactory way to live with the consequences of their diabetes (DFT $\leq$4) with patients who still perceived a burden of diabetes on their daily functioning (DFT >4).

The secondary outcome measure 'diabetes-related emotional distress' was measured by the 20-item Problem Areas in Diabetes questionnaire (PAID).[29–31] Scores were transformed to a 0–100 scale, with 0 being the most favourable outcome. Missing items were imputed using patients' individual mean score if at least 50% of items were available.

Participation and autonomy were measured by means of the Impact on Participation and Autonomy (IPA) questionnaire.[32] The response options of the 32 items ranged from 0 to 4 with higher scores representing poorer participation and autonomy. Five subscale scores were computed: autonomy indoors (range 0–28), family role (range 0–28), autonomy outdoors (range 0–20), social relationships (range 0–28) and work and education (range 0–24). In accordance with the manual, missing items were imputed using the individual mean scores of items that were not missing in patients for whom at least 75% of items were available.[33]

Patients' self-management knowledge and behaviours regarding their diabetes were measured using the Dutch version of the Partners in Health scale (PIH-NL).[34] The 12 items covered four domains of patients' competency

in relation to their self-management (knowledge, coping, management of condition and adherence to treatment). We used the total sum score.

The 12-item Short-Form Health Survey (SF-12) measured the quality of life.[35] A physical component summary and a mental component summary were computed by using item weights and regression constants for the Dutch population derived from the oblique rotation method.[36 37] The norm-based summary scores have a mean of 50 and an SD of 10 in the Dutch population.[37] Higher scores reflect better outcomes on the quality of life.

The General Self-Efficacy Scale (GSES-12) assessed patients' belief in their ability to organise and engage in certain behaviours.[38 39] A higher sum score (range 12–60) reflects higher levels of self-efficacy. Missing items were imputed using patients' individual mean score if at least 50% of items were available.

The glycated haemoglobin in mmol/mol was measured during consultations. For the baseline measure, we used data from the first consultation after the patient's baseline measurement (with a maximum interval of 3 months). The consultation following this first consultation was used as the 4-month follow-up measurement. The last measurement was the assessment approximately 1 year after the patient's baseline measurement (at least more than 6 months after the first consultation).

Statistical analysis

The power calculation was based on the dichotomous DFT.[16] On the basis of a group size of 46 PNs, we determined in advance that a sample size of 232 patients (at least 5 patients per PN) would have 90% power and an α of 0.05 to detect an improvement in perceived daily functioning (defined as DFT ≤4) at 12 months measurement occurring in 20% of patients in the intervention arm versus 5% of those in the control arm. In this power calculation for cluster-randomised trials, an intraclass correlation coefficient of 0.04 was used. Assuming that not all positively screened patients would give informed consent for trial participation, and a 30% loss to follow-up, we planned to invite 10 eligible patients for each PN.

Baseline variables were compared by means of independent-samples t tests and χ^2 tests to detect statistically significant differences between both groups at baseline. For the effect evaluation, linear and logistic multilevel models were used with an unstructured covariance structure for repeated measures, and patients and PNs as random factors. These models account for the correlation between the repeated measurements for each patient and also for the hierarchical structure of the data, with measurements (level-1) clustered within patients (level-2) who in turn were clustered within PNs (level-3). Type of practice, that is, the stratification variable, time of measurement (categories: 0, 4 and 12) and the interaction between group and time of measurement were included in the models as fixed effects. Analyses were performed on an intention-to-treat basis.

To ensure that all participants of the intervention and control arms started with a score DFT >4, indicating a perceived burden of diabetes, we used the dichotomised research-led screening outcomes at baseline. Consequently, the model only included 4-month and 12-month follow-ups as the baseline value was the same for all participants.

A two-sided p value <0.05 was considered statistically significant. Statistical analyses were conducted using the software packages IBM SPSS Statistics V.21.0 for Windows and SAS V.9.3.

RESULTS
Participants

From the 77 family practices that were approached between April and June 2011, 40 agreed to participate. Their PNs (n=41) were by randomisation assigned to the intervention arm (20 PNs) and the control arm (21 PNs). After randomisation, but before patient recruitment, one family practice in the intervention arm withdrew from study participation due to the heavy workload of the PN. This left 19 PNs who received training in SMS and integrated it into their daily practice, and 21 PNs in the control arm who provided usual care.[12] As some PNs worked together in a team, 15 units of analysis were left in the intervention arm and 19 units of analysis in the control arm.

A total number of 3822 patients with diabetes were sent a screening questionnaire in the period from October 2011 to March 2012. The response rate was 50% in the intervention arm and 44% in the control arm. As screening questionnaires were sent by GPs, we could not compare responders with non-responders with regard to demographic characteristics.

From all 1805 patients who participated in the screening, 357 (20%) met the detection criteria and were invited to take part in the trial. A total of 264 patients (74%) gave informed consent to participate in the trial. Their research-led screening outcomes were not significantly different from outcomes from eligible patients who refused trial participation.

From the 117 intervention patients and 147 control patients who participated in the trial, 80% completed all of the follow-up measurements. In 10% of the sample, one follow-up measurement was missing. Three patients did not complete the baseline measurement and gave informed consent at the 4-month follow-up measurement. Another 23 patients completed only the baseline measurement. We found no baseline variables that were significantly related to incompleteness of measurements. Figure 2 illustrates the flow of recruitment and randomisation of PNs, screening of patients and measurements of trial participants.

Baseline characteristics

Table 1 illustrates the baseline characteristics of participants in the SMS trial.

The sample comprised 46% female patients with the majority having a lower level of education. The average age of the participants was 65. Of all 121 participants younger than 65 years, 41% had paid employment. The proportion of patients with a paid job was significantly higher in the intervention arm than in the control arm. On average, patients had been diagnosed with diabetes for 8–9 years. The majority were being treated with oral hypoglycaemic medication. The number of patients being treated with insulin was significantly higher in the intervention group (33%) compared to patients in the control group (16%). At the time of baseline measurement, 84% were not receiving psychosocial care from a health professional.

Patients of both groups were comparable for the primary and secondary outcomes at the baseline measurement except for the sum score on the PIH scale. This reflected better self-management skills of intervention patients regarding knowledge, symptom management and adherence to treatment. Both groups were comparable regarding the subdomain 'coping' that included three items about dealing with the effects of diabetes on daily life.

Study participants' exposure to SMS

An unexpected outcome was the very low exposure to SMS of the study participants in the intervention arm. It is presented in table 2. PNs did not register any data about SMS

Figure 2 Flow of participants through the Self-Management Support (SMS) trial. DFT, Daily Functioning Thermometer; PNs, practice nurses.
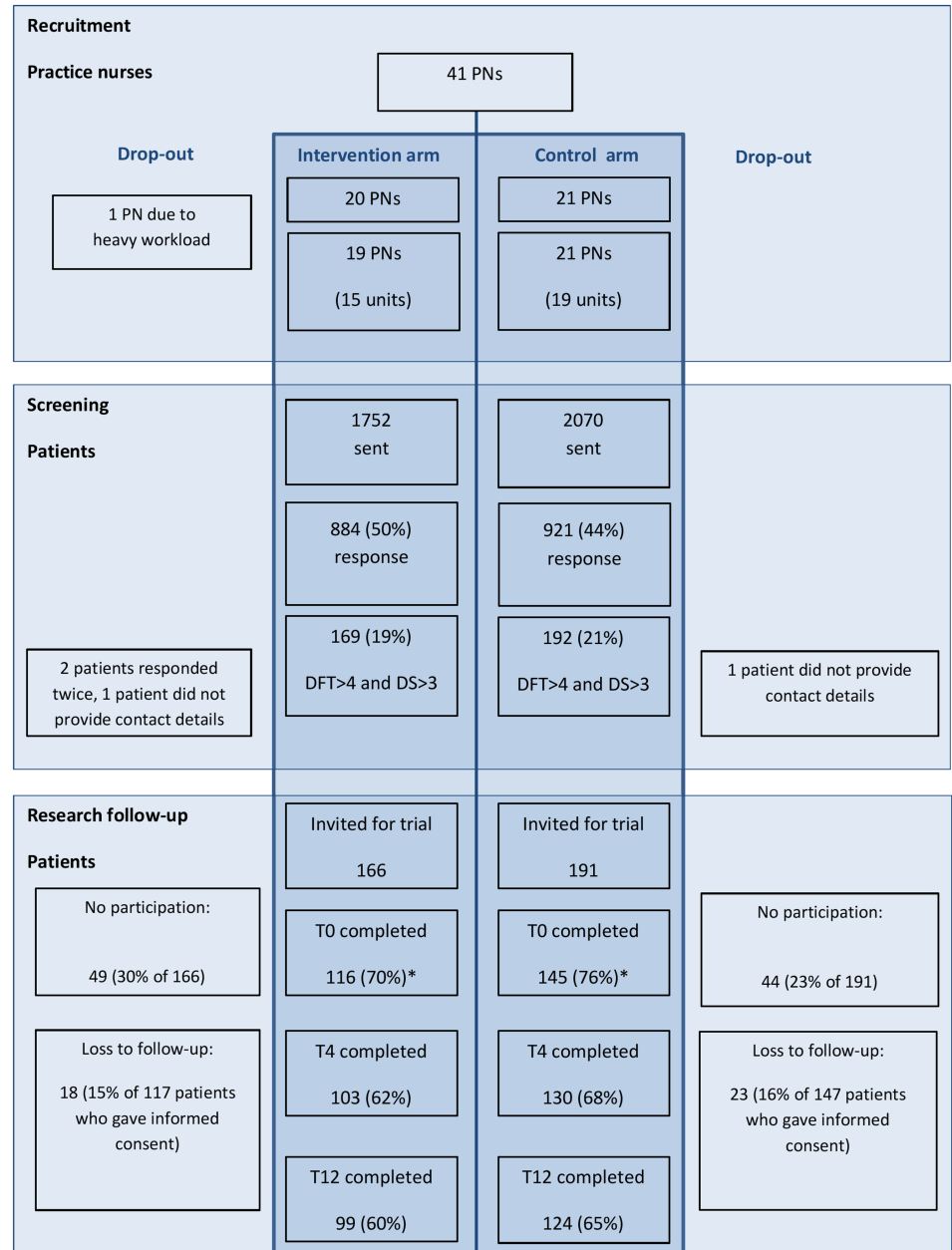
*Some patients (n=1 in the intervention arm and n=2 in the control arm) did not respond at baseline, but agreed to participate in the trial at time of the T4 measurement.

for 17 participants (15%) in the patient file during the year of follow-up. For the other 100 participants, between 1 and 5 diabetes consultations were registered with data on the DFT and DS. The combination of DFT >4 and DS >3 was found in only 16 participants (14%).

A total of 46 study participants scored neither DFT >4 nor DS >3 during consultations. The low detection rates also implied low exposure to the follow-up phase of SMS.

From 30 participants who scored DS >3 and were therefore eligible for further diagnostics, 10 patients did not receive the 4DSQ because (1) they were already undergoing psychological treatment, or (2) they were not interested, or (3) a specific stressor was clearly causing the distress. Extra consultations for SMS were provided for 11 participants, of whom there were only two patients with a nurse-led detection score of DFT >4 and DS >3. From the other 14 participants with positive scores for DFT and DS, six patients were referred to the GP and one patient to a mental nurse because severe mental health problems were recorded using the 4DSQ. Five participants with score DFT >4 and DS >3 did not want to receive follow-up care, and two participants were already receiving psychological treatment.

Primary outcome
As illustrated in table 3, the number of patients who improved in their daily functioning (score DFT ≤4) was lower in the intervention group than in the control

Table 1 Baseline characteristics of study participants in intervention and control arms

Characteristics	Categories	Intervention arm (n=117)	Control arm (n=147)	
Gender	Female	55 (47)	67 (46)	
Age	Mean (SD) age (years)	64 (10)	65 (9)	
Diagnosis of diabetes	Mean (SD) duration (years)	9 (8)	8 (6)	
Ethnicity	Non-western	2 (1.7)	0 (0)	
Education*	Low	80 (72)	103 (74)	
Work status of patients <65 years†	Paid job	28 of 55 (51)	21 of 66 (32)	
Marital status	Married	76 (66)	92 (64)	
	Single	5 (4)	12 (8)	
	Divorced	16 (14)	21 (14)	
	Widowed	19 (16)	19 (13)	
Treatment†	Diet only	8 (7)	12 (8)	
	Tablets	69 (61)	108 (76)	
	Insulin	10 (9)	4 (3)	
	Insulin and tablets	27 (24)	19 (13)	
Psychological care	No psychological care	96 (83)	123 (86)	
	In primary care setting	17 (15)	16 (11)	
	In secondary care setting	3 (3)	4 (3)	
Health outcomes	Scale (range)	Mean (SD)	Mean (SD)	p Value t Test
Daily functioning	DFT (0–10) at screening	6.8 (1.3)	6.6 (1.3)	0.205
Blood glucose	HbA1c in mmol/mol	53.0 (11.2)	51.8 (10.2)	0.429
Diabetes-related distress	PAID (0–100)	29.9 (16.9)	28.9 (19.4)	0.684
Participation and autonomy	IPA autonomy indoors (0–28)	6.9 (4.8)	6.7 (4.5)	0.734
	IPA family role (0–28)	12.7 (5.5)	12.8 (5.7)	0.921
	IPA autonomy outdoors (0–20)	8.6 (4.2)	8.4 (4.2)	0.681
	IPA social relationships (0–28)	9.6 (4.6)	9.4 (4.3)	0.631
	IPA work and education (0–24)	11.1 (3.5)	9.4 (3.2)	0.116
Self-management	PIH (0–96)	78.0 (8.9)	73.2 (14.5)	0.002†
Quality of life	SF-12 Physical component	34.8 (9.6)	35.0 (9.8)	0.849
	SF-12 Mental component	34.1 (11.3)	35.2 (11.2)	0.456
Self-efficacy	GSES (12–60)	38.6 (7.5)	39.2 (7.0)	0.481

Values are numbers (percentages) unless stated otherwise.
*Low refers to primary school, lower vocational training or lower general education.
†Significant differences: p<0.05.
HbA1c, glycated haemoglobin. DFT, Daily Functioning Thermometer; PAID, Problem Areas in Diabetes; IPA, Impact on Participation and Autonomy; lower mean scores reflect better outcomes. PIH, Partners in Health Scale; SF-12, Short Form Health Questionnaire Physical and Mental component, mean score Dutch population is 50; GSES, General Self-efficacy scale; higher scores refer to better outcomes.

group at 4-month follow-up, but this was not a statistically significant difference (24% vs 32%; OR: 0.505 (95% CI 0.213 to 1.201)). At the 12-month follow-up, more patients in the intervention group improved in their daily functioning compared to patients in the control group. However, this was not statistically significant either (33% vs 26%, OR 1.754 (95% CI 0.742 to 4.148)).

Secondary outcomes

For the secondary outcome measures, the intervention group perceived significantly more participation than the control group regarding autonomy indoors and family role (p<0.05) after 4 months. The same applied, with nearly statistically significant differences, for autonomy outdoors (p=0.06) and social relationships (p=0.05). However, these effects disappeared at 12-month follow-up. As illustrated in table 4, we found

no other statistically significant differences for the secondary outcome measures at 4 and 12 months.

DISCUSSION

This study, with a low exposure of study participants to the complete intervention, could not demonstrate any effect of SMS on emotional, social or biomedical parameters, except for autonomy and participation with regard to activities indoors and the family role after 4 months. This effect in favour of the patients in the intervention arm had disappeared after 12 months.

The critical issue of the implementation of SMS was the addition of a detection procedure as an inherent and integrated part of SMS. PNs were trained to identify eligible patients for the follow-up phase of SMS by applying screening questions during diabetes consultations.

Table 2 Exposure to SMS of study participants in the intervention arm (n=117)

	N	Outcome detection	4DSQ completed of those who received the 4DSQ*	4DSQ Outcome	Self-Management Support	Referral
	46	DFT ≤4 and DS ≤3	–	–	2	–
			1 of 2	1 mild	–	–
	24	Only DFT >4	–	–	1	
			3 of 3	2 mild	–	–
				1 moderate	1	
	14	Only DS >3	5 of 8	2 moderate	2	
				3 severe	3	
	16	DFT >4 and DS >3	12 of 12	4 moderate	2	1
				8 severe	–	6
	17	No registration	–	–	–	–
Total	117		21 of 25		11	7

*PNs were instructed to give the 4DSQ to patients with score DS >3.
HbA1c, glycated haemoglobin. DFT, Daily Functioning Thermometer; PAID, Problem Areas in Diabetes; IPA, Impact on Participation and Autonomy; lower mean scores reflect better outcomes. PIH, Partners in Health Scale; SF-12, Short Form Health Questionnaire Physical and Mental component, mean score Dutch population is 50; GSES, General Self-efficacy scale; higher scores refer to better outcomes.

The DFT and DS were considered to be simple, patient-centred indicators of patients' daily functioning and their emotional distress. The 4DSQ would enable the PNs and GPs to get more insight into the presence and severity of the emotional problems. For the effectiveness trial, a self-administered postal screening procedure was inserted into the study design to select patients from both the intervention and control arms similarly. To perform a pragmatic trial, we minimised the interference of the researchers in clinical practice as much as possible. The outcomes of DFT and DS on the research-driven screening were consistent with our assumptions based on the literature that approximately 20% of the diabetic population would be eligible for the follow-up phase of SMS. However, these efforts to select patients with emotional problems, aimed at increasing the contrast between patients in the intervention and control arms, appeared to be useless as the added detection method in routine practice did not function as expected. This was not a problem of PNs' adherence to the SMS protocol, as the screening tools were integrated into diabetes consultations of more than 85% of the study participants. The problem was that the majority of study participants who scored above the cut-off values on the postal questionnaires did not meet the required detection criteria when screened face to face by the PNs. These patients were therefore not exposed to the follow-up phase of SMS. We have no reason to assume that non-respondents on the research-driven screening were those most in distress, who would have been the easiest to be detected by the PNs, as low detection rates were reported for the whole diabetic population in the intervention practices.

Differences between the self-administered screening outcomes and the nurse-led detection could be explained by the fluctuation in distress symptoms and the phenomenon of regression towards the mean. This may have resulted in study participants scoring less extreme values at nurse-led detection compared to values at research-driven screening. Furthermore, tests may function differently across settings and administration methods.[40] Owing to patients' difficulty in expressing emotional problems during consultations and lack of recognition by health professionals of emotional problems experienced by patients with chronic physical illness, the chance of positive screening outcomes seems to have been higher in anonymous research-driven screening than in clinical practice.[41 42] Nonetheless, we did not expect such a large discrepancy between research-driven screening and nurse-led detection.

For 17 of the 117 study participants, no data on SMS activities were registered, suggesting that these patients were exposed neither to the detection phase nor the follow-up phase of SMS. The reasons for this phenomenon could be twofold. First, PNs had to integrate SMS into all consultations while the effects were assessed in only a few patients per PN. PNs did not know which

Table 3 Multilevel analyses for differences between intervention and control arms regarding improvement on patients' daily functioning reflected by the dichotomous primary outcome (DFT ≤4) at 4-month follow-up and at 12-month follow-up

	Intervention	Control	OR	95% CI	p Value
DFT ≤4	Yes/N (%)	Yes/N (%)			
4 months	24/102 (23.5)	40/126 (31.7)	0.505	0.213 to 1.201	0.12
12 months	32/96 (33.3)	32/121 (26.4)	1.754	0.742 to 4.148	0.20

DFT, Daily Functioning Thermometer.

Table 4 Multilevel analyses for differences between intervention arm and control arm for secondary outcome measures at 4-month and 12-month follow-ups

Outcome measure (range)	Scale (range)	Mean (SD)		Adjusted treatment effect			
		Intervention arm	Control arm	Difference in mean	95% CI		p Value
At 4-month follow-up							
Blood glucose	HbA1c in mmol/mol	54.7 (11.0)	52.1 (9.6)	−0.03	−1.97	1.91	0.98
Diabetes-related distress	PAID (0–100)	26.1 (16.5)	27.0 (19.7)	−2.22	−5.46	1.01	0.18
Participation and autonomy	IPA indoors (0–28)	5.9 (4.4)	7.0 (5.1)	−1.27	−2.25	−0.30	0.01*
	IPA family role (0–28)	11.4 (5.3)	12.8 (6.0)	−1.25	−2.33	−0.17	0.02*
	IPA outdoors (0–20)	7.5 (4.0)	8.1 (4.3)	−0.69	−1.43	0.03	0.06
	IPA social relationships (0–28)	8.4 (4.3)	9.3 (4.7)	−0.83	−1.64	−0.01	0.05
	IPA work and education (0–24)	10.3 (5.7)	9.4 (5.7)	+0.01	−2.14	2.17	0.99
Self-management	PIH (0–96)	79.9 (7.9)	76.3 (12.3)	+0.01	−2.19	2.21	0.99
Quality of life	SF-12 Physical component	35.4 (10.0)	35.7 (10.4)	+0.17	−1.62	1.96	0.85
Quality of life	SF-12 Mental component	37.1 (11.4)	38.5 (11.7)	−0.99	−3.54	1.55	0.44
Self-efficacy	GSES (12–60)	39.5 (7.1)	39.8 (7.4)	+0.15	−1.05	1.36	0.81
At 12-month follow-up							
Blood glucose	HbA1c in mmol/mol	51.5 (10.8)	50.4 (9.6)	−0.80	−2.75	1.15	0.42
Diabetes-related distress	PAID (0–100)	27.8 (17.7)	24.2 (16.9)	+0.65	−2.64	3.93	0.70
Participation and autonomy	IPA indoors (0–28)	6.6 (4.6)	7.4 (5.2)	−0.63	−1.63	0.36	0.21
	IPA family role (0–28)	12.0 (6.2)	12.5 (5.9)	−0.13	−1.24	0.97	0.81
	IPA outdoors (0–20)	8.5 (4.1)	7.9 (4.0)	+0.56	−0.19	1.30	0.14
	IPA social relationships (0–28)	9.1 (4.4)	8.9 (4.0)	−0.01	−0.84	0.81	0.98
	IPA work and education (0–24)	11.4 (5.2)	11.5 (4.5)	+0.02	−2.09	2.13	0.98
Self-management	PIH (0–96)	78.6 (8.6)	77.9 (10.7)	−2.25	−4.49	−0.01	0.05
Quality of life	SF-12 Physical component	36.3 (10.5)	34.9 (10.6)	+0.33	−1.48	2.14	0.72
Quality of life	SF-12 Mental component	37.5 (11.9)	37.3 (10.8)	+0.43	−2.15	3.01	0.74
Self-efficacy	GSES (12–60)	38.6 (7.6)	40.3 (6.9)	−0.69	−1.92	0.54	0.27

*Statistically significant difference (p<0.05).
DFT, Daily Functioning Thermometer; GSES, General Self-efficacy scale; higher scores refer to better outcomes; HbA1c, glycated haemoglobin; PAID, Problem Areas in Diabetes; IPA, Impact on Participation and Autonomy; lower mean scores reflect better outcomes; PIH, Partners in Health Scale. SF-12=Short Form Health Questionnaire Physical and Mental component, mean score Dutch population is 50.

patients participated in the trial. As PNs' integration of SMS into consultations could have fluctuated during follow-up, they may have missed study participants. Second, registration bias may have occurred. PNs needed to open an extra data file to record the process and outcomes of SMS. This step might have created a barrier to their compliance. This underlines the need for a thorough process evaluation of how and under what conditions the intervention procedure was carried out. In this regard, the simultaneous evaluation of the implementation process will reveal important issues. The outcomes of this process evaluation will be presented elsewhere.

In the planning of the study, we considered avoiding research-driven screening by inserting the nurse-led detection phase of SMS into both study arms. However, monitoring and discussing psychological well-being as part of routine care may already have positive effects on patients' moods.[43] Moreover, from an ethical perspective, patients in the control arm who would be detected positively for mental health problems would need to be referred for further diagnosis or psychosocial treatment,

thus changing the usual care. Using research-led screening was therefore assumed to be the most adequate solution for the selection of participants for the effectiveness part of the study. In retrospect, a pilot study should have been part of our implementation project. It was not planned for several reasons. There was an implementation moment. Besides, we could use the experiences from the former study, in which the treatment phase was evaluated. PNs would receive comprehensive training, and they would ask common questions about psychosocial functioning. Furthermore, the outcomes of the screening did not give rise to questions about the instruments: as expected, 20% of the respondents met the detection criteria.

Although collaboration with regional stakeholders resulted in organisational and financial benefits for family practices willing to participate, this incentive did not appear to be a decisive factor in the process of considering whether or not to participate. For robust multilevel modelling, it is necessary that sufficient clusters are recruited and sufficient patients are available per

cluster.[44] Our intention was to include 46 PNs, but only 40 PNs participated. Furthermore, the final total number of clusters for analysis was actually only 33 as some patients were seen alternately by PNs working in a team. Owing to the regional approach of the SMS project, we were confined to the family practices within the region under study. The number of patients within clusters ranged from 1 to 19. The loss to follow-up of patients was less than expected, which led to an acceptable total number of study participants. However, the availability of clusters as well as the balance between clusters needs attention when designing a cluster-randomised trial that will be performed in an implementation setting.

In implementation studies, formative evaluation methods can be used to adapt or optimise the implementation or intervention itself during the study.[20] Owing to the trial protocol and time and research funding limitations, we did not further optimise the intervention or implementation strategies other than to inform PNs after 3-month follow-up to use the criteria DFT >4 and DS >3 less rigidly. The planning of a more flexible design of the effectiveness trial within the hybrid design would have been helpful in overcoming implementation problems. This calls for inclusion of specific evaluation moments to decide about the need for adjustments in the intervention itself, in implementation activities or with regard to follow-up measurements.

In conclusion, the adjusted nurse-led intervention in its present form was not effective. There was a dilution of the contrast between intervention and control patients. This is because outcomes of the detection method, added for implementation purposes, were neither optimal nor congruent with outcomes of the screening method added for research purposes. We do not know whether SMS would be effective if PNs used other methods to identify patients in need for support in their emotional and role management skills. Blending effectiveness and implementation studies may have the potential to facilitate a rapid adoption of evidence-based care innovations in routine practice and to provide useful information for decision-makers about effects and implementation issues,[20] but researchers who are planning a hybrid effectiveness-implementation study need to be aware of the necessity to build in adequate follow-up moments and to take action either to remove identified barriers or to adapt the intervention to an acceptable level.

Author affiliations
[1]Department of Family Medicine, Maastricht University, CAPHRI School for Public Health and Primary Care, Maastricht, The Netherlands
[2]Department of Social Medicine, Maastricht University, CAPHRI School for Public Health and Primary Care, Maastricht, The Netherlands
[3]Department of Methodology and Statistics, Maastricht University, CAPHRI School for Public Health and Primary Care, Maastricht, The Netherlands
[4]Department of General Practice and Elderly Care Medicine, VU University Medical Centre, EMGO Institute for Health and Care Research, Amsterdam, The Netherlands

REFERENCES

1. Fisher EB, Thorpe CT, DeVellis BM, *et al*. Healthy coping, negative emotions, and diabetes management. *Diabetes Educ* 2007;33:1080–103.
2. Holman H, Lorig K. Patient self-management: a key to effectiveness and efficiency in care of chronic disease. *Public Health Rep* 2004;119:239–43.
3. Penninx BWJH, Kritchevsky SB, Yaffe K, *et al*. Inflammatory markers and depressed mood in older persons: results from the health, aging and body composition study. *Biol Psychiatry* 2003;54:566–72.
4. Fisher L, Mullan JT, Arean P, *et al*. Diabetes distress but not clinical depression or depressive symptoms is associated with glycemic control in both cross-sectional and longitudinal analyses. *Diabetes Care* 2010;33:23–8.
5. DiMatteo MR, Lepper HS, Croghan TW. Depression is a risk factor for noncompliance with medical treatment: meta-analysis of the effects of anxiety and depression on patient adherence. *Arch Int Med* 2000;160:2101–7.
6. Aikens JE. Prospective associations between emotional distress and poor outcomes in type 2 diabetes. *Diabetes Care* 2012;35:2472–8.
7. Macdonald W, Rogers A, Blakeman T, *et al*. Practice nurses and the facilitation of self-management in primary care. *J Adv Nurs* 2008;62:191–9.
8. Van Houtum L, Rijken M, Heijmans M, *et al*. Self-management support needs of patients with chronic illness: do needs for support differ according to the course of illness? *Patient Educ Couns* 2013;93:626–32.
9. Rane K, Wajngot A, Wändell PE, *et al*. Psychosocial problems in patients with newly diagnosed diabetes: number and characteristics. *Diabetes Res Clin Pract* 2011;93:371–8.
10. Elissen A, Nolte E, Knai C, *et al*. Is Europe putting theory into practice? A qualitative study of the level of self-management support

in chronic care management approaches. *BMC Health Serv Res* 2013;13:117.

11. Netherlands Diabetes Federation. *NDF Care Standard. Transparancy and quality of diabetes care for people with type 2 diabetes*. Amersfoort: Nederlandse Diabetes Federatie (NDF), 2007.

12. Rutten GEHM, De Grauw WJC, Nijpels G, *et al*. NHG-Standaard Diabetes mellitus type 2 (Tweede herziening) [NHG Practice Guideline Diabetes mellitus type 2 (Second revision)]. *Huisarts Wet* 2006;49:137–52.

13. Tsiachristas A, Hipple-Walters B, Lemmens KMM, *et al*. Towards integrated care for chronic conditions: Dutch policy developments to overcome the (financial) barriers. *Health Policy* 2010;101:122–32.

14. Jansen D, Spreeuwenberg P, Heijmans M. *Ontwikkelingen in de zorg voor chronisch zieken: rapportage 2012 [Developments in chronic care: report 2012]*. Utrecht: Netherlands Institute for Health Services Research (NIVEL), 2012.

15. Heiligers PJM, Noordman J, Korevaar JC, *et al. Kennisvraag: praktijkondersteuners in de huisartspraktijk (POH's) klaar voor de toekomst? [Knowledge base: Practice nurses in the GP practices, ready for the future?]*. Utrecht: Netherlands Institute for Health Services Research (NIVEL), 2012.

16. Van Dijk-de Vries A, van Bokhoven MA, Terluin B, *et al*. Integrating nurse-led Self-Management Support (SMS) in routine primary care: design of a hybrid effectiveness-implementation study among type 2 diabetes patients with problems of daily functioning and emotional distress: a study protocol. *BMC Fam Pract* 2013;14:77.

17. Lamers F, Jonkers CCM, Bosma H, *et al*. A minimal psychological intervention in chronically ill elderly patients with depression: a randomised trial. *Psychother Psychosom* 2010;79:217–26.

18. Jonkers CCM, Lamers F, Bosma H, *et al*. The effectiveness of a minimal psychological intervention on self-management beliefs and behaviors in depressed chronically ill elderly persons: a randomized trial. *Int Psychogeriatr* 2012;24:288–97.

19. Lamers F, Jonkers CCM, Bosma H, *et al*. Treating depression in diabetes patients: does a minimal psychological intervention affect diabetes-specific quality of life and glycemic control? A randomized controlled trial. *J Adv Nur* 2011;67:788–99.

20. Curran GM, Bauer M, Mittman B, *et al*. Effectiveness-implementation hybrid designs: combining elements of clinical effectiveness and implementation research to enhance public health impact. *Med Care* 2012;50:217–26.

21. Tuinman MA, Gazendam-Donofrio SM, Hoekstra-Weebers JE. Screening and referral for psychosocial distress in oncologic practice. *Cancer* 2008;113:870–8.

22. Maas-van Schaaijk NM, Odink RJ, Ultee K, *et al*. Can one question be a useful indicator of psychosocial problems in adolescents with diabetes mellitus? *Acta Paediatrica* 2011;100:708–11.

23. Braam C, van Oostrom SH, Terluin B, *et al*. Validation study of a distress screener. *J Occup Rehabil* 2009;19:231–7.

24. Struijs JN, Baan CA. Integrating care through bundled payments—lessons from the Netherlands. *N Engl J Med* 2011;364:990–1.

25. Terluin B, van Marwijk HWJ, Adèr HJ, *et al*. The Four-Dimensional Symptom Questionnaire (4DSQ): a validation study of a multidimensional self-report questionnaire to assess distress, depression, anxiety and somatization. *BMC Psychiatry* 2006;6:34.

26. Van Eijk JTM, Diederiks JPM, Kempen GIJM, *et al*. Development and feasibility of a nurse administered strategy on depression in community-dwelling patients with a chronic physical disease. *Patient Educ Couns* 2004;54:87–94.

27. Mynors-Wallis L. Problem solving treatment in general psychiatric practice. *Advan Psychiatr Treat* 2001;7:417–25.

28. Lamers F, Jonkers CCM, Bosma H, *et al*. Effectiveness and cost-effectiveness of a minimal psychological intervention to reduce non-severe depression in chronically ill elderly patients: the design of a randomised controlled trial. *BMC Public Health* 2006;6:161.

29. Polonsky WH, Anderson BJ, Lohrer PA, *et al*. Assessment of diabetes-related distress. *Diabetes Care* 1995;18:754–60.

30. Snoek FJ, Pouwer F, Welch GW, *et al*. Diabetes-related emotional distress in Dutch and US diabetic patients: cross-cultural validity of the problem areas in diabetes scale. *Diabetes Care* 2000;23:1305–9.

31. Welch G, Weinger K, Anderson B, *et al*. Responsiveness of the Problem Areas In Diabetes (PAID) questionnaire. *Diabet Med* 2003;20:69–72.

32. Cardol M, de Haan RJ, de Jong BA, *et al*. Psychometric properties of the Impact on Participation and Autonomy Questionnaire. *Arch Phys Med Rehabil* 2001;82:210–16.

33. Kersten P. Impact on Participation and Autonomy (IPA) Manual to the English version: IPA. http://www.nivel.nl/sites/default/files/bestanden/INT-IPA-Manual.pdf (accessed 18 Jun 2015).

34. Petkov J, Harvey P, Battersby M. The internal consistency and construct validity of the partners in health scale: validation of a patient rated chronic condition self-management measure. *Qual Life Res* 2010;19:1079–85.

35. Ware JE, Kosinski M, Keller S. A 12-item short-form health survey: construction of scales and preliminary tests of reliability and validity. *Med Care* 1996;34:220–33.

36. Fleishman JA, Selim AJ, Kazis LE. Deriving SF-12v2 physical and mental health summary scores: a comparison of different scoring algorithms. *Qual Life Res* 2010;19:231–41.

37. Mols F, Pelle AJ, Kupper N. Normative data of the SF-12 health survey with validation using postmyocardial infarction patients in the Dutch population. *Qual Life Res* 2009;18:403–14.

38. Bosscher RJ, Smit JH. Confirmatory factor analysis of the General Self-Efficacy Scale. *Behav Res Ther* 1998;36:339–43.

39. Sherer M, Maddux JE, Mercandante B, *et al*. The self-efficacy scale: construction and validation. *Psychol Rep* 1982;51:663–71.

40. De Waal MW, van der Weele GM, van der Mast RC, *et al*. The influence of the administration method on scores of the 15-item Geriatric Depression Scale in old age. *Psychiatry Res* 2012;197:280–4.

41. Cape J, McCulloch Y. Patients' reasons for not presenting emotional problems in general practice consultations. *Br J Gen Pract* 1999;49:875–9.

42. Mitchell AJ, Rao S, Vaze A. Can general practitioners identify people with distress and mild depression? A meta-analysis of clinical accuracy. *J Affect Disord* 2011;130:26–36.

43. Pouwer F, Snoek FJ, van der Ploeg HM, *et al*. Monitoring of psychological well-being in outpatients with diabetes. *Diabetes Care* 2001;24:1929–35.

44. Campbell MK, Mollison J, Steen N, *et al*. Analysis of cluster randomized trials in primary care: a practical approach. *Fam Pract* 2000;17:192–6.