An evaluation of the effectiveness and cost effectiveness of the National Exercise Referral Scheme in Wales, UK: a randomised controlled trial of a public health policy initiative

Simon Mark Murphy,[1] Rhiannon Tudor Edwards,[2] Nefyn Williams,[3] Larry Raisanen,[1] Graham Moore,[1] Pat Linck,[2] Natalia Hounsome,[2] Nafees Ud Din,[3] Laurence Moore[1]

[1]DECIPHer, Cardiff School of Social Sciences, Cardiff University, Cardiff, Wales, UK
[2]Centre for Health Economics and Medicines Evaluation, Institute of Medical and Social Care Research, Bangor University, Bangor, UK
[3]North Wales Centre for Primary Care Research, Bangor University, North Wales Clinical School, Wrexham, UK

Correspondence to
Dr Simon Mark Murphy, DECIPHer, Cardiff School of Social Sciences, Cardiff University, Cardiff, Wales, UK; murphys7@cf.ac.ukre

Accepted 26 March 2012
Published Online First 10 May 2012

ABSTRACT

Background The Wales National Exercise Referral Scheme (NERS) is a 16-week programme including motivational interviewing, goal setting and relapse prevention.

Method A pragmatic randomised controlled trial with nested economic evaluation of 2160 inactive participants with coronary heart disease risk (CHD, 1559, 72%), mild to moderate depression, anxiety or stress (79, 4%) or both (522, 24%) randomised to receive (1) NERS or (2) normal care and brief written information. Outcome measures at 12 months included the 7-day physical activity recall, the hospital anxiety and depression scale.

Results Ordinal regression identified increased physical activity among those randomised to NERS compared with those receiving normal care in all participants (OR 1.19, 95% CI 0.99 to 1.43), and among those referred for CHD only (OR 1.29, 95% CI 1.04 to 1.60). For those referred for mental health reason alone, or in combination with CHD, there were significantly lower levels of anxiety (OR −1.56, 95% CI −2.75 to −0.38) and depression (OR −1.39, 95% CI −2.60 to −0.18), but no effect on physical activity. The base-case incremental cost-effectiveness ratio was £12 111 per quality adjusted life year, falling to £9741 if participants were to contribute £2 per session.

Conclusions NERS was effective in increasing physical activity among those referred for CHD risk only. Among mental health referrals, NERS did not influence physical activity but was associated with reduced anxiety and depression. Effects were dependent on adherence. NERS is likely to be cost effective with respect to prevailing payer thresholds.

Trial registration Current Controlled Trials ISRCTN47680448.

INTRODUCTION

It is widely recognised that regular physical activity is beneficial to both physical and mental health[1] It is associated with reduced risk from chronic diseases, including coronary heart disease (CHD)[2] and has been shown to be positively linked to mental health,[3] including depression.[4] Exercise referral schemes (ERS) can target specific patient or population subgroups with such conditions by providing contact with qualified exercise professionals (EP) and access to tailored programmes promoting physical activity.

Despite the rapid growth of ERS in the UK, the evidence base for their effectiveness and cost effectiveness is equivocal. The latest systematic review evidence prior to this study identified six randomised controlled trials from the UK,[5] where a modest but statistically significant improvement in activity with a combined RR of 1.2 (95% CI 1.06 to 1.35) was partly explained by poor rates of uptake and adherence, and a lack of intervention relapse-prevention strategies. Overall, results were consistent with previous reviews in that ERS increased physical activity in individuals who were already slightly active,[6 7] increases were not maintained long-term[7] and scheme attendance was poor.[8] International cost-effectiveness evidence was also equivocal,[5 9] with a pre-study review[10] identifying nine studies varying from €348 to €86 877 (£304 to £75 982) per quality adjusted life year (QALY) depending on scheme intensity.

In light of this, and the development of variable localised ERS throughout the UK, rigorous evidence is needed to distinguish scheme content that facilitates uptake and adherence and promotes long-term improvements in activity.[11] This paper reports an independent evaluation of the Welsh Government's National Exercise Referral Scheme (NERS) operating in 12 local health board (LHB) areas in Wales, UK, assessing its effectiveness and cost effectiveness in increasing physical activity and reducing anxiety and depression among patients referred for CHD risk and/or anxiety, depression and stress.

METHODS

Study design

A pragmatic randomised controlled trial, with nested process and economic evaluations (for full details, see published study protocol[12]).

Recruitment of participants

Those eligible for the scheme were sedentary and had at least one medical condition (table 1). Patients were identified opportunistically by clinicians in normal practice and were provided with basic trial information by the clinician who completed a referral form forwarded to the evaluation team. Those referred for reasons other than CHD and mental health could access the scheme outside of the trial, while those referred with CHD risk factors and/or mild to moderate anxiety,

Table 1 Scheme inclusion/exclusion criteria used by clinicians in normal practice and trial eligibility

Scheme inclusion criteria	Scheme exclusion criteria	Trial eligibility
The patient must be **sedentary** (defined as not moderately active for ≥3 times per week or deconditioned through age or inactivity), and have at least one of the following medical conditions: ► CHD risk factors – Raised blood pressure more than 140/90 (either) but <180/100 (either) – Weight management – BMI >28 – Controlled diabetes – Impaired glucose tolerance – High cholesterol >5.0 – Family history of heart disease or diabetes – Referral from Cardiac Rehabilitation Schemes (only from phase IV) ► Mental health – Mild anxiety, depression or stress ► Musculoskeletal – At risk of osteoporosis – Arthritis (mild) – Poor mobility – Musculoskeletal pain including back pain ► Respiratory/pulmonary – COPD – Mild/moderate well controlled – (asthma, bronchitis, emphysema) ► Neurological conditions – Multiple sclerosis ► Other – Smoker – Chronic fatigue	► Aged ≤16 years ► Unstable angina ► Blood pressure 180/100 (in either) or above and/or uncontrolled or poorly controlled hypertension ► Cardiomyopathy ► Uncontrolled tachycardia ► Cardiac arrhythmia ► Valvular heart disease ► Congenital heart disease ► Unexplained dizzy spells ► Excessive or unexplained breathlessness on exertion ► Uncontrolled or poorly controlled diabetes ► Uncontrolled or poorly controlled epilepsy ► History of falls or dizzy spells in the last 12 months ► Uncontrolled or poorly controlled asthma (severe COPD) ► First 12 weeks of pregnancy ► Awaiting medical investigation ► Aneurysms ► History of cerebro-vascular disease ► Unstable/newly diagnosed angina (within 6 months) ► Established coronary heart disease (including myocardial infarction) ► Any other uncontrolled condition	The patient must be **sedentary** and have at least one of the following condition: ► CHD risk factors – raised blood pressure more than 140/90 (either) but <180/100 (either) – weight management – BMI >28 – controlled diabetes – impaired glucose tolerance – high cholesterol >5.0 – family history of heart disease or diabetes – referral from Cardiac Rehabilitation Schemes (only from phase IV), and or ► Mental health – mild anxiety, depression or stress

BMI, body mass index; CHD, coronary heart disease; COPD, chronic obstructive pulmonary disease.

depression or stress were eligible for the trial and were sent full informed consent materials and a brief baseline questionnaire, with non-responders sent a repeat mailing 2 weeks later. The postal questionnaire requested demographic details, post code of residence and the general practice physical activity questionnaire (GPPAQ).[13] Eligible patients declining to participate in the trial entered a 12-month waiting list.

Randomisation

Each participant who consented and returned a completed baseline questionnaire was assigned a unique ID and entered sequentially into the study database. These were randomly assigned to the intervention ERS or control trial arm using a random number generator, with gender and LHB as stratification variables. Randomisation of forwarded referral forms occurred every 2 weeks, with treatment allocation blind and remote from participants and practitioners.

Intervention

The intervention followed a standardised protocol and was delivered at leisure centres by exercise professionals (EP) in each LHB[14 15] (box 1). The content, duration and intensity of the scheme were designed to promote adherence and long-term improvements in activity. Consultations were based on motivational interview[16] principles which facilitated patient-centred achievable goals, and included relapse-prevention strategies at 4 and 16 weeks to review goals and encourage attendance.[14] The primary goal was for participants to achieve 30 min of moderate physical activity on at least 5 days per week. EPs delivering the programme were not directly aware of whether or not a client was a trial participant but could potentially identify this on the basis of the reason for referral. Blinding of participants was not feasible. The control group received usual care and a leaflet highlighting the benefits of exercise, and were given the addresses of local facilities.

Sample size

Sample size was determined to detect a difference in total minutes of weekly activity at 12 months, with 1052 participants in each group providing 90% power to detect an effect size of 0.15 with no loss to follow-up and, more realistically, 87% and 84% power to detect an effect size of 0.15 if 20% and 25%, respectively, of randomised participants who were lost to follow-up.

Box 1 Delivery of the Welsh National Exercise Referral Scheme (NERS)

16-week tailored programme of exercise supervised by a qualified exercise professional

- Initial consultation with exercise professional on entry: lifestyle questionnaire, health check (resting heart rate, blood pressure, body mass index and waist circumference), introduction to leisure centre facilities, motivational interview and goal setting.
- Access to one-to-one exercise instruction and/or group exercise classes. Discounted rate for exercise activities, £1 per session.
- Four-week telephone contact with exercise professional—review of goals, motivational interview, relapse prevention.
- Sixteen-week consultation with exercise professional—review of goals, motivational interview, health check, lifestyle questionnaire, service evaluation questionnaire[14] and advice on continuing with exercise after the programme.

Post-16-week activities

- 8-months telephone contact by exercise professional to ask about their exercise behaviour and relapse prevention.
- 12-months review including repeat of health check carried out at entry and Chester fitness step test.[15]

Outcome measures

The primary outcome was total minutes of weekly physical activity at 12-month follow-up, assessed using the 7-day physical activity recall (7D-PAR)[17] administered by telephone[18] with interviewees blind to group allocation. For those telephone respondents unwilling to complete the 7D-PAR, the briefer GPPAQ measure was administered where possible.

A postal questionnaire was also sent to all study participants at 6 and 12 months, with non-responders sent a repeat mailing 2 weeks later. At 6 and 12 months, participants completed an adapted Client Service Receipt Inventory,[19] the health-related quality of life measure EQ-5D[20] and willingness-to-pay question.[21] The outcome measure for the economic analysis was the QALY derived using the EQ-5D, a generic preference-based instrument for measuring health-related quality of life.[20] The Client Service Receipt Inventory questionnaire[19] asked participants to recall their contacts with NHS primary care (including prescribing) and secondary care services over the preceding 6 months (baseline) and at 6 and 12 months after baseline. The questionnaire at 12 months also included a question asking participants how much (in UK pounds) they were, in theory, willing to pay for exercise sessions through NERS.[21] At 12 months, the questionnaire included the Baecke Questionnaire of Habitual Physical Activity (Baecke)[22] and the Hospital Anxiety and Depression Scale (HADS) (1983)[23] to assess depression and anxiety as secondary outcomes.

Analysis

As the primary outcome (7-D PAR) has a highly skewed and bimodal distribution, it was recoded according to approximate quintiles as a five-level ordinal variable, and proportional odds ordinal regression models were used with the stratification variables (gender, LHB area, age group (16–44, 45–59, 60+)) and baseline activity level (GPPAQ) as covariates. Secondary analysis excluded baseline activity level as a covariate. Analyses were repeated with imputed values for those who did not complete the 7-D PAR, but who did complete either the Baecke or GPPAQ at 12 months, using stochastic imputation based on their Baecke or GPPAQ measures. Subgroup analyses for gender, age group (16–44, 45–59, 60+), referral reason (mental health only, CHD only, or combination of CHD and mental health), and tertile of Welsh Index of Multiple Deprivation obtained for the lower-layer super-output area of the participants' postcode of residence,[24] was assessed by including the main effect and interaction in separate models. We used the same approach with linear regression to explore our secondary outcome measures (HADS) for mental health referrals, or those referred for mental health/CHD combined. The analyses of HADS among all participants are secondary to this analysis. Analyses were conducted on an intention-to-treat basis with the statistician unaware of how the treatment group variable was coded. For each outcome per protocol, analyses to identify whether outcomes vary in terms of adherence to the programme replaced the binary intervention variable with a three-level programme attendance variable; full attendance (for 16 weeks), partial attendance (1–16 weeks), and no exposure to programme (control group or non-attender).

Economic evaluation

Costs and benefits of NERS were estimated from a public sector perspective and were not discounted, as follow-up was for 1 year. A primary cost-utility analysis was conducted using the base-case intervention cost per participant (£385; n=3530). As

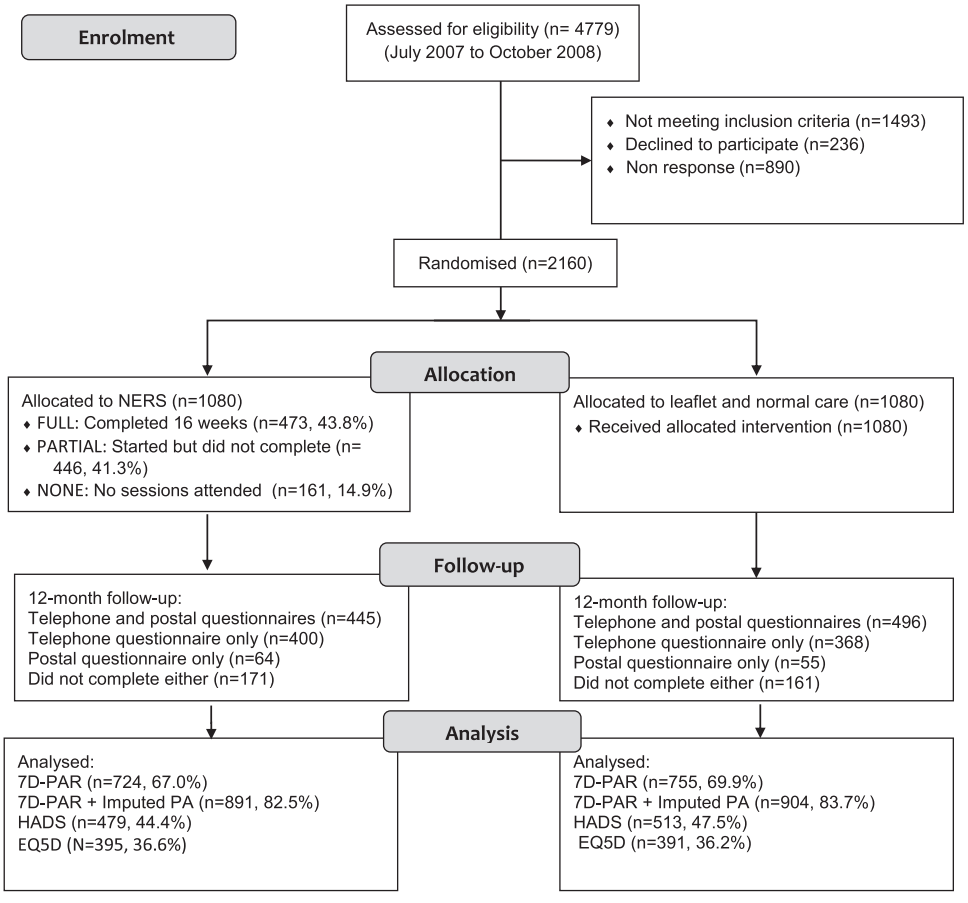

Figure 1 Flow diagram of the study. EQ5D, EuroQol—5 Dimensions; HADS, Hospital Anxiety and Depression Scale; NERS, National Exercise Referral Scheme; 7-D PAR, 7-day physical activity recall.

EQ-5D was not included in the minimal data collection at baseline in this trial, conservatively, 6-month EQ-5D values were used as a baseline estimate to generate an incremental cost-effectiveness ratio (ICER) at 12 months and cost effectiveness acceptability curve to compare with the National Institute of Health and Clinical Excellence (NICE) cost per QALY threshold of £20 000–£30 000.[25] National unit costs were applied to service use frequency data to estimate total costs of service use for patients,[26–28] and cost per QALY estimates were calculated using utility weights from EQ-5D.[29] When EQ-5D data were missing for 1 or 2 domains in the 5-domain scale (n=26), stochastic imputation was used. A participant payment of either £1 or £2 per session (based on the findings from our willingness-to-pay analysis) was included in the sensitivity analysis, with a mean attendance of 2 sessions per week for the full 16-week programme at either £32 or £64. Economic subgroup analysis was conducted for reason of referral, age group, gender and adherence.[30] Cost-effectiveness analysis was conducted using Stata SE version 10. A nonparametric Mann–Whitney U test was used to compare HR-QoL due to the skewedness of the EQ-5D data.

RESULTS

Participant flow and follow-up

Figure 1 shows participant flow through the study and numbers available for analysis; 1479 (68.5%) for 7D-PAR, 1795 (83.1%) for imputed 7D-PAR and 992 (45.9%) for HADS. Response rates were similar in the two groups. Of those allocated to the intervention, 43.8% (n=473) completed the 16-week programme, 41.3% (n=446) started the programme but did not complete it and 14.9% (n=161) failed to attend. Table 2 shows baseline characteristics for those completing 7D-PAR (n=1479) and HADS measures (n=992) at 12-month follow-up. Although there is some evidence of greater loss to follow-up among younger participants and those referred in whole or part for mental health reasons, there was no strong evidence of differential loss to follow-up in terms of gender or deprivation.

Baseline data

Participants were aged between 16 and 88 years (mean 52, SD 14.7), predominately women (66%) and the vast majority classed themselves as white (96%). Table 2 shows participants were most likely to be referred for CHD risk factors only (72%) or in combination with mental health issues (24%) and classed themselves as inactive (58.6%) or moderately inactive (15.3%), with 24% defining themselves as either active or moderately active. The economic analysis was based on 55% (n=798) of the participants in the effectiveness analysis. The economic sample contained fewer younger participants (n=140, 18%) than the main trial (n=1423, 30%), and included a higher proportion of participants who were referred for CHD risk factors only (n=616, 77%) compared with the main sample (n=1559, 71%). More of them also completed the 16-week programme (62%, n=247), with fewer partial (32%, n=123) and non-attenders (8%, n=30). The intervention and control groups were similar for all baseline characteristics.

Table 2 Comparison of demographic characteristics by trial arm at baseline and for EuroQol—5 Dimensions (EQ5D) and 12-month outcomes

	Baseline		EQ5D		7-D PAR		HADS	
	Intervention % (N)	Control % (N)	Intervention % (N)	Control % (N)	Intervention % (N)	Control % (N)	Intervention % (N)	Control % (N)
Reasons for referral								
CHD only	71.3 (770)	73.1 (789)	76.8 (307)	77.6 (309)	74 (536)	75.8 (572)	75.2 (360)	76.6 (393)
Mental health only	3.8 (41)	3.5 (38)	3.3 (13)	3.3 (13)	3.5 (25)	3.1 (23)	2.9 (14)	2.9 (15)
CHD and mental health	24.9 (269)	23.4 (253)	20.0 (80)	19.1 (976)	22.5 (163)	21.2 (160)	21.9 (105)	20.5 (105)
Age group								
16–44	29.9 (322)	30.1 (323)	17.3 (69)	17.8 (71)	26.6 (192)	26.0 (196)	18.8 (90)	19.5 (100)
45–59	34.5 (371)	32.7 (352)	33.8 (135)	34.0 (136)	35.3 (255)	34.3 (258)	34.7 (166)	35.1 (180)
60+	35.6 (383)	37.2 (400)	49.0 (196)	48.0 (191)	38.1 (275)	39.7 (299)	46.4 (222)	45.4 (233)
Gender								
Male	34.4 (372)	34.5 (373)	33.8 (135)	32.7 (130)	34.5 (250)	33.3 (251)	35.1 (168)	33.1 (170)
Female	65.6 (708)	65.5 (707)	66.3 (265)	67.0 (268)	65.5 (474)	66.7 (504)	64.9 (311)	66.9 (343)
Welsh index of multiple deprivation tertile								
Low	34.4 (361)	32.3 (340)	35.0 (140)	32.2 (128)	37.1 (262)	33.5 (246)	37.5 (173)	32.9 (165)
Middle	34.1 (358)	32.5 (342)	33.5 (134)	38.0 (152)	32.4 (232)	35.0 (257)	34.3 (158)	37.3 (187)
High	31.4 (330)	35.2 (370)	27.8 (111)	27.9 (111)	30 (212)	31.6 (232)	28.2 (130)	29.7 (149)
General practice physical activity questionnaire								
Inactive	59.2 (623)	60.6 (643)	56.3 (225)	60.8 (242)	58.2 (411)	61.4 (455)	59.0 (276)	60.5 (306)
Moderate inactive	16.1 (170)	15.1 (160)	14.8 (59)	13.6 (54)	15.6 (116)	15.3 (113)	14.3 (67)	14.0 (71)
Moderate active	17.2 (181)	15.2 (161)	17.8 (71)	14.6 (58)	19.3 (136)	14.8 (110)	17.5 (82)	15.2 (77)
Active	7.5 (79)	9.2 (97)	8.5 (34)	9.8 (39)	6.4 (49)	8.5 (63)	9.2 (43)	10.3 (52)
Missing	(27)	(19)	2.8 (11)	1.3 (5)				
Employment								
Employed	32.7 (346)	28.4 (300)	26.3 (105)	24.6 (98)	32.7 (232)	29.5 (218)	27.0 (131)	27.7 (139)
Retired	30 (318)	32.9 (348)	41.3 (165)	41.7 (166)	32.7 (232)	35.1 (259)	38.8 (183)	40.2 (202)
Housework	18.8 (199)	20.3 (214)	16.3 (65)	18.9 (75)	17.0 (121)	20.6 (152)	17.0 (80)	18.5 (93)
Other	18.5 (196)	18.5 (195)	16.3 (62)	13.1 (52)	17.6 (125)	14.8 (109)	16.5 (78)	13.6 (68)
Missing	1.9 (21)	2.1 (23)	0.8 (3)	2.0 (8)				
Education								
Beyond min school leaving age	52.1 (557)	53.0 (570)	57.0 (228)	58 (231)	51.5 (370)	54.6 (410)	58.7 (280)	55.5 (284)
Total (n)	1080	1080	400	398	724	755	479	513

CHD, coronary heart disease; 7-D PAR, 7-day physical activity recall.

Intervention effects

Table 3 presents the 12-month median scores and inter-quartile range for the 7-D PAR and means and CIs for depression and anxiety scores by trial arm and age, gender, Welsh Index of Multiple Deprivation, adherence, reason for referral and baseline GPPAQ. Table 4 shows the results of the regression analyses for each of the primary outcomes at 12-month follow-up. For all participants, those in the intervention group had higher levels of physical activity than those in the control, but this was of borderline statistical significance. Among those referred for CHD risk factors, the intervention group reported significantly higher levels of activity, but there was no difference among those referred wholly or partially for mental health reasons. Among this group of referrals, those randomised to NERS had significantly lower levels of both depression and anxiety.

Subgroup analyses showed effectiveness was highly dependent on adherence, with significantly greater differences in all outcomes among those who completed the 16-week programme compared with those who attended only partially or not at all. There were also significant interactions with gender for both mental health outcomes, with the beneficial effect of the intervention only apparent among women. There was a suggestion that the intervention was more effective on mental health outcomes among the youngest age group (18–44), although this was not statistically significant. Effects did not vary significantly by deprivation status.

Cost effectiveness

The data on health-related quality of life and adherence to the programme are summarised in table 5. A significant difference in HR-QoL between the intervention and control groups was found using EQ-5D-VAS. For participants <44 years of age, the difference between both EQ-5D and VAS scores was significant. 62% (n=247) of the sample upon which economic analysis was undertaken completed the 16-week programme, 32% (n=123) attended fewer than 16 weeks and 8% (n=30) did not attend at all. There were no significant differences in NHS resource use between the intervention and control groups, except that the control groups were referred for significantly more health-related tests (p=<0.05) (data not presented). In the base-case analysis, the difference in costs between intervention and control group was £327, and the difference in QALYs was 0.027, which generated an ICER point estimate of £12 111 per QALY gained. The probability of the intervention being cost effective was 89% at the NICE threshold of £30 000 per QALY.

The results of sensitivity and subgroup analyses are summarised in table 6. Using the mean intervention and control group 6-month EQ-5D score as an estimate for the baseline, QALYs resulted in a decrease of the ICER from £12 111 (base case) to £6055 per QALY. When only the control group's mean EQ-5D value at 6 months was used as an estimate of baseline QALYs for both control and intervention groups, the ICER point estimate was £7109. This analysis demonstrates that using the 6-month EQ-5D data as an estimate of baseline, QALYs in the base-case analysis was a rather conservative approach. When possible, participant payments of £1 and £2 per session were added to the base-case analysis, the cost per QALY fell to £10 926 and £9741, respectively. Subgroup analyses found that the intervention is likely to be more cost effective in: participants with CHD and/or mental health risk factors compared with participants with a risk of CHD only; female rather than male participants; younger (<44 years) rather than older individuals. Subgroup analysis based on those who had adhered fully to the

Table 3 Twelve-month follow-up: medians and IQR for 7-day PAR and means and 95% confidence intervals for HADS by trial arm and covariates

		7-day PAR total minutes of exercise median (IQR)		HADS anxiety score mean (95% CI)		HADS depression score mean (95% CI)	
Co-variate	**Level**	**Intervention**	**Control**	**Intervention**	**Control**	**Intervention**	**Control**
All		200 (60, 435) n=724	165 (50, 370) n=755	7.82 (7.39 to 8.25) n=472	8.35 (7.92 to 8.77) n=502	6.14 (5.73 to 6.54) n=471	6.93 (6.53 to 7.32) n=506
Age	16–44	210 (62.5, 452.5) n=192	207.5 (60, 425) n=196	9.05 (8.08 to 10.07) n=88	10.80 (9.97 to 11.63) n=95	7.39 (6.39 to 8.40) n=89	8.70 (7.87 to 9.54) n=98
	45–59	210 (60, 520) n=255	127.5 (35, 355) n=258	9.39 (8.64 to 10.13) n=165	9.13 (8.39 to 9.87) n=177	7.10 (6.38 to 7.81) n=164	7.81 (7.09 to 8.52) n=177
	60+	185 (50, 390) n=275	175 (60, 355) n=299	6.14 (5.57 to 6.70) n=218	6.73 (6.16 to 7.30) n=230	4.91 (4.40 to 5.43) n=217	5.50 (4.98 to 6.01) n=231
Gender	Female	200 (60, 455) n=474	150 (42.5, 345) n=504	7.64 (7.10 to 8.17) n=308	8.83 (8.32 to 9.33) n=333	5.70 (5.21 to 6.19) n=305	7.22 (6.74 to 7.69) n=337
	Male	205 (75, 395) n=250	210 (60, 420) n=251	8.16 (7.43 to 8.89) n=164	7.40 (6.65 to 8.16) n=169	6.94 (6.25 to 7.63) n=166	6.35 (5.65 to 7.05) n=169
WIMD	Low	205 (80, 400) n=262	165 (35, 360) n=246	7.19 (6.46 to 7.91) n=171	7.29 (6.56 to 8.02) n=162	5.43 (4.83 to 6.03) n=167	6.28 (5.59 to 6.97) n=164
	Middle	205 (60, 557.5) n=232	160 (60, 405) n=257	7.85 (7.13 to 8.56) n=156	8.28 (7.59 to 8.98) n=183	6.30 (5.60 to 7.00) n=157	6.54 (5.91 to 7.16) n=183
	High	170 (45, 390) n=212	172.5 (60, 392.5) n=232	8.53 (7.65 to 9.41) n=127	9.49 (8.69 to 10.29) n=145	6.81 (5.96 to 7.67) n=129	8.25 (7.49 to 9.01) n=147
Adherence	0 wk.	140 (20, 370) n= 79	NA	9.17 (7.73 to 10.61) n=47	NA	7.36 (6.05 to 8.66) n=45	NA
	> 0 to <16 wk. adherence	162.5 (37.5, 385) n=284	NA	8.63 (7.86 to 9.40) n=159	NA	7.14 (6.39 to 7.89) n=158	NA
	16 wk.	240 (90, 510) n=361	NA	7.09 (6.55 to 7.64) n=266	NA	5.34 (4.85 to 5.84) n=268	NA
Reason for referral	CHD only	210 (80, 442.5) n=536	170 (60, 360) n=572	6.95 (6.49 to 7.41) n=355	7.33 (6.88 to 7.78) n=386	5.45 (5.03 to 5.87) n=355	6.18 (5.76 to 6.60) n=389
	Mental health only	220 (30, 315) n=25	240 (60, 450) n=23	12.85 (9.96 to 15.73) n=13	12.07 (9.55 to 14.58) n=15	9.69 (6.77 to 12.62) n=13	9.27 (6.34 to 12.19) n=15
	CHD and mental health reasons	155 (40, 420) n=163	147.5 (32.5, 392.5) n=160	10.15 (9.25 to 11.06) n=104	11.68 (10.81 to 12.55) n=101	8.06 (7.09 to 9.03) n=103	9.42 (8.51 to 10.33) n=102

HADS, Hospital Anxiety and Depression Scale; WIMD, Welsh index of multiple deprivation; 7-D PAR, 7-day physical activity recall.

Table 4 Main effects and interaction effects - ORs and 95% CIs from ordinal logistic regression models examining impacts of NERS on physical activity and B coefficients and 95% CIs from linear regression models examining impacts of NERS on depression and anxiety

Effect	Group	7-D PAR	7-D PAR plus imputed values	HADS depression	HADS anxiety
Main effects	Whole sample				
	All covariates included (n=1443/1749/959/956)†	1.19 (0.99, 1.43)	1.18 (0.99, 1.42)	−0.71* (−1.25, −0.17)	−0.54 (−1.12, 0.35)
	Baseline GPPAQ omitted (n=1475/1788/976/973)†	1.20* (1.00, 1.45)	1.19* (1.00, 1.42)	−0.74* (−1.28, −0.20)	−0.49 (−1.06, 0.08)
	CHD only				
	All covariates included (n=1081/1302/732/729)†	1.29* (1.04, 1.60)	1.26* (1.02, 1.57)	−0.60* (−1.18, −0.02)	−0.32 (−0.95, 0.31)
	Baseline GPPAQ omitted (n=1105/1329/743/740)†	1.35** (1.09, 1.67)	1.30* (1.05, 1.60)	−0.64* (−1.22, −0.03)	−0.27 (−0.90, 0.35)
	MH only or MH and CHD				
Interaction effects	All covariates included (n=362/447/227/227)†	1.06 (0.73, 1.55)	1.04 (0.72, 1.52)	−1.39* (−2.60, −0.18)	−1.56* (−2.75, −0.38)
	Baseline GPPAQ omitted (n=370/459/233/233)†	0.94 (0.65, 1.36)	1.00 (0.69, 1.36)	−1.32* (−2.54, −0.10)	−1.52* (−2.68, −0.37)
	Deprivation (n=1405/1698/929/926)†				
	Medium x intervention	1.08 (0.69 to 1.70)	0.98 (0.63 to 1.55)	0.80 (−0.50 to 2.09)	−0.24 (−1.64 to 1.15)
	High x intervention	0.83 (0.52 to 1.32)	0.80 (0.52 to 1.25)	−0.13 (−1.50 to 1.25)	−0.46 (−1.94 to 1.02)
	Adherence level (n=1443/1749/959/956)†				
	Partial	1.00 (0.78 to 1.29)	1.00 (0.79 to 1.27)	−0.12 (−0.90 to 0.65)	−0.12 (−0.84 to 0.82)
	Full	1.46* (1.17 to 1.84)	1.40* (1.11 to 1.79)	−1.24* (−1.88 to −0.61)	−1.12* (−1.80 to −0.44)
	Gender (n=1443/1749/959/956)†				
	Male x intervention	0.75 (0.51 to 1.10)	0.74 (0.51 to 1.07)	2.10* (0.98 to 3.23)	1.93* (0.72 to 3.14)
	Age (n=1443/1749/959/956)†				
	45−59 x intervention	1.36 (0.84 to 2.21)	1.36 (0.84 to 2.18)	0.61 (−0.94 to 2.16)	2.03* (0.37 to 3.69)
	60+ x intervention	0.99 (0.63 to 1.58)	1.04 (0.67 to 1.62)	0.65 (−0.83 to 2.13)	1.07 (−0.52 to 2.66)

*p<0.05; **p<0.01.
†n represents number of patients in 7-day PAR analyses/7-day PAR plus imputed values/HADS depression/HADS anxiety. All models include gender, LHB area, age group as covariates. Except where stated, all models also include baseline GPPAQ as a further covariate.
CHD, coronary heart disease; GPPAQ, general practice physical activity questionnaire; HADS, Hospital Anxiety and Depression Scale; MH, mental health; 7-D PAR, 7-day physical activity recall.

16-week programme indicated a saving of −£367 per QALY gained.

DISCUSSION

Among those referred for CHD risk factors only, the NERS in Wales was associated with significantly higher levels of physical activity when compared with normal care. However, among those referred for mental health reasons, either solely or in combination with CHD, there was no difference in physical activity between the NERS and normal care participants at 12-month follow-up. The primary analysis of the trial was of the impact of the scheme on all referrals, and this was of borderline statistical significance, being in effect a pooled estimate of two heterogeneous subgroup effects. Two recent systematic reviews[31 32] have highlighted the need to examine variation in effectiveness based on medical condition, a view strongly supported by the current study.

Table 5 Mean cost-effectiveness outcomes by group over 12 months

Outcomes at 12 month follow-up	Intervention group mean (SD)	Control group mean (SD)	*p value
EQ-5D score (0−1)			
Entire sample	0.64 (0.32) n=395	0.61 (0.32) n=391	0.09
<44 years	0.68 (0.33) n=69	0.54 (0.34) n=71	0.012*
45−60 years	0.58 (0.35) n=132	0.59 (0.35) n=131	0.88
>60 years	0.67 (0.28) n=194	0.65 (0.27) n=89	0.37
EQ-5D VAS (0−100)			
Entire sample	68.8 (18.6) n=397	63.9 (20) n=394	0.0007*
<44 years	67.62 (19.1) n=69	58.3 (20.2) n=71	0.003*
45−60 years	66.1 (18.6) n=135	62.0 (19.5) n=133	0.13
>60 years	71.1 (18.2) n=193	67.3 (19.8) n=190	0.07
Adherence			
Did not attend	30 (8%)	NA	NA
Attended <16 weeks	123 (32%)	NA	NA
Attended 16 weeks	247 (62%)	NA	NA

*Significant (p<0.05) according to Mann−Whitney test.
EQ5D, EuroQol—5 Dimensions; VAS, visual analogue scale.

For patients referred for mental health reasons, the scheme was not effective in helping them to increase physical activity. Further planned analyses of data collected at 6 months will allow an assessment of whether the scheme was effective in increasing self-efficacy and motivation to exercise among these patients. Consistent with recent systematic review findings,[31 32] patients referred for mental health reasons did appear to benefit in terms of reduced anxiety and depression, particularly in this trial among women and younger patients. This suggests that for these patients, the EP's attention and the social contact and support generated by scheme attendance may be the beneficial mechanisms rather than increased physical activity per se, a hypothesis requiring further research.

One of the recent reviews also highlights the importance of scheme uptake, adherence and their predictors in explaining outcomes: NERS uptake at 85% was slightly above the average (80%) found in the review.[32] Although only 44% of intervention participants adhered to the scheme throughout, this compares favourably with the pooled rate of 37% across schemes assessed by trials in the review.[32] The relationship between adherence, psychosocial processes and 12-month outcomes will be assessed in future papers.

A pre-study review suggested that low-contact and low-intensity physical activity interventions are more cost effective than more intensive interventions.[10] NERS is an intensive 16-week programme which we have found to be likely to be cost effective at conventional thresholds, a finding consistent with recent cost-effectiveness reviews,[32 33] which modelled lifetime benefits and showed a 51% probability of cost effectiveness at £20 000 per QALY and 88% at £30 000 per QALY.

Strengths and weaknesses of the study

A pre-study systematic review of ERS found only six RCTs in the UK and identified a number of shortcomings, including

Table 6 Cost-effectiveness sensitivity (A) and subgroup analysis (B)

	Intervention group (n)	Control group (n)	Cost of intervention (£)	ICER (£)	Bootstrapped one-sided 95% CI for ICER (£)	Probability that intervention is cost effective*
(A) Sensitivity analyses						
Base case	395	391	385	12 111	58 881	89%
Mean intervention and control EQ-5D at 6 months as estimate of baseline	395	391	385	6055	37 159	96%
Control group mean EQ-5D value at 6 months as estimate of baseline	395	391	385	7109	24 853	98%
Participant payment of £1 per session	395	391	385−32=353	10 926	69 085	91%
Participant payment of £2 per session	395	391	385−64=321	9741	64 638	92%

	Intervention group (n)	Control group (n)	Incremental cost (bootstrapped 95%CI)	Incremental QALY (bootstrapped 95%CI)	ICER (bootstrapped one-sided 95%CI)
(B) Subgroup analysis					
Referral reason					
CHD only	307	309	£239 (−51 to 547)	0.0183 (−50.0047 to 0.0416)	£13 060 (117 893)
Mental health (MH) and MH plus CHD	80	76	£596 (−5304 to 1616)	0.058 (0.017 to 0.100)	£10 276 (50 925)
Gender					
Male	135	130	£322 (−5180 to 792)	0.0084 (−0.0250 to 0.0400)	£38 333 (254 973)
Female	265	268	£326 (−589 to 761)	0.0362 (0.0111 to 0.0623)	£9006 (390 00)
Age					
<44 years	69	71	£68 (−5462 to 678)	0.0656 (0.0137 to 0.1148)	£1037 (16 418)
44−60 years	135	136	£577 (−5176 to 1373)	0.0179 (−0.0194 to 0.0561)	£32 235 (314 108)
>60 years	196	191	£244 (−5131 to 586)	0.0187 (−0.0057 to 0.0424)	£13 048 (153 565)
Adherence					
0 weeks	30		£1785 (25 to 4570)	−0.0114 (−0.0680 to 0.0390)	NA
<16 weeks	123		£662 (212 to 1157)	−0.0084 (−0.0384 to 0.0218)	NA
16 weeks	247		−£18 (−5310 to 277)	0.049 (0.0277 to 0.0706)	−£367 (7068)

*At the £30 000 per QALY threshold.
EQ5D, EuroQol—5 Dimensions; ICER, incremental cost-effectiveness ratio; QUALY, quality adjusted life year.

blinding of outcome measurement, the need for long-term follow-up and the generalisability of the study population.[5] A more recent review of eight European RCTs[31] suggested that substantial heterogeneity in the quality and nature of schemes may have contributed to inconsistent evidence of their effectiveness and identified the need for further high-quality RCTs of theoretically informed approaches to behaviour change (including motivational interviewing). They also highlighted the need to assess subgroup effectiveness.

In NERS, motivational interviewing was used as a clearly identified approach to behaviour change, though implementation checks indicated that it was often poorly delivered[16] with data indicating that in practice, the key active ingredients of the programme were the professionals' support and supervision and interaction with other patients.[34] This suggests that scheme effectiveness could be improved with increased attention to fidelity of motivational interviewing. The importance of subgroup analysis based on medical condition is highlighted in the inconsistent outcomes for CHD and mental health referrals observed in the current study. In addition, the primary outcome was collected at 12 months by researchers blinded to condition. Physiological outcomes and objective measures of physical activity were not collected. However, the 7D-PAR is well validated and was administered by telephone which provides a more reliable measure than postal questionnaires.[35] Confidence in the robustness of findings is enhanced by the relatively high response rate achieved for the primary outcome (68.5%) compared with equivalent community trials of ERS.[5] Using other measures of physical activity (Baecke and GPPAQ) to impute scores for weekly activity for those without a 7D PAR measure provided an enhanced response rate of 83.1%, although it should be acknowledged that GPPAQ was developed as a clinical screening, rather than a data-capture tool.

Significantly, as a pragmatic policy evaluation, results are likely to have high external validity and generalisabilty.[36] There was minimal control over the implementation of the intervention, other than randomisation, and it was delivered across Wales by a wide range of professionals to patients from areas covering the full spectrum of socioeconomic circumstances. The NERS in Wales shares many features of similar schemes implemented across the UK.[37]

It should also be noted that there was a much lower response rate to the 12-month postal questionnaire, and thus the economic and mental health analyses are based on a smaller number of participants. Although there was no strong patterning in response rates, these may be subject to unmeasured response bias, and it is acknowledged that cost effectiveness is assessed at 12 months only.

As in previous studies,[5–7] it should be noted that a significant minority of trial participants, who were referred on the basis of their clinician identifying them as sedentary, reported activity levels at baseline which classed them as active or moderately active by GPPAQ criteria. The inclusion of this subgroup may

What is already known on this subject

- Previous evaluations of Exercise Referral Schemess in the UK have found only modest improvements in physical activity in the short term, and there is uncertainty whether such approaches are cost effective. These effects are partly explained by poor rates of uptake and adherence to the schemes and a lack of intervention relapse-prevention strategies.

What this study adds

- At 12-month follow-up, the Wales NERS was associated with increased physical activity among those referred for coronary heart disease risk factors. Among those referred for mental health reasons, there were reductions in depression and anxiety without any increase in physical activity.
- Programme intensity and the provision of relapse strategies appear to be effective in promoting relatively high rates of adherence to the scheme, which was associated with greater improvements on all outcomes.

Policy implications

- Applying the NICE threshold of £20 000—£30 000 per QALY to wider government spending, the Wales National exercise Referral Scheme has a moderate to high probability of being cost effective.
- Pragmatic effectiveness trials nested within government policy roll-outs are feasible and provide opportunities to develop generalisable public health intervention research evidence.

well have contributed to higher levels of adherence and activity, however, as a pragmatic policy trial it was important that these individuals were not excluded. Given the number of policy and practice constraints that have inhibited such evaluations,[11] the current study provides a relatively rare example of a pragmatic randomised trial of a national ERS and demonstrates the feasibility of calls for an increase in such trials of public health interventions.[38]

CONCLUSION

NERS was effective in increasing physical activity among those referred with CHD risk factors. Although there was no increase in physical activity among those referred for mental health reasons, anxiety and depression were reduced. These effects were highly dependent on adherence to the programme. NERS is likely to be cost effective under prevailing payer thresholds.

Acknowledgements The authors thank the scheme's coordinators, Elaine McNish and Jeannie Wyatt-Williams and Janine Hale and Chris Roberts of the Welsh government for providing information and ongoing support to facilitate this study. The authors are grateful to all exercise coordinators and professionals and patients that took part in the study. The research was independent and funded by the Welsh Assembly Government. Additional support during writeup was provided by The Centre for the Development and Evaluation of Complex Interventions for Public Health Improvement (DECIPHer), a UKCRC Public Health Research: Centre of Excellence. Funding from the British Heart Foundation, Cancer Research UK, Economic and Social Research Council (RES-590-28-0005), Medical Research Council, the Welsh Assembly Government and the Wellcome Trust (WT087640MA), under the auspices of the UK Clinical Research Collaboration, is gratefully acknowledged. The Thames Valley Multi-centre Research Ethics Committee (MREC) approved the evaluation of the Welsh NERS on 8 Feb 2007 (Ref: 06/MRE12/85). Due to changes in COREC procedures, an additional Site-Specific Information form (Ref: C/92139/137055/1) was also later submitted to the MREC. Approval from medical directors within each local health board has also been obtained in partnership with local exercise coordinators. The conduct of the study conformed to the principles embodied in the Declaration of Helsinki.

Contributors Principal responsibility for the study was assumed by SMM. LR was the trial manager and responsible for the day-to-day running of the study. GM was responsible for the development and day-to-day running of the mixed-method process evaluation. RTE was responsible for study design and overseeing of the economic evaluation. PL was responsible for the day-to-day running of the economic evaluation. NH was responsible for the cost-effectiveness analysis. NW contributed to the study design and aspects of the process evaluation. NUD was responsible for the design and conduct of aspects of the process evaluation. LM contributed to the study design, conducted sample size calculations, designed the analysis plan and took responsibility for statistical analysis. SM and RTE produced a first draft of the manuscript, and SM was responsible for the final revised version, developing and integrating contributions. All authors read and commented on drafts and approved the final manuscript and had full access to all the data (including statistical reports and tables) in the study and can take responsibility for the integrity of the data and the accuracy of the data analysis. SMM acts as guarantor.

Funding This study was supported by the Welsh Government.

Competing interests None.

Ethics approval Ethics approval was provided by the Thames Valley Multi-centre Research Ethics Committee (MREC) approved the evaluation of the Welsh NERS on 8 Feb 2007 (Ref: 06/MRE12/85).

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement Data is still currently being analysed to address secondary research questions.

REFERENCES

1. **Woodcock J,** Franco O, Orsini N, *et al*. Non-vigorous physical activity and all-cause mortality: systematic review and meta-analysis of cohort studies. *Int J Epidemiol* 2010;**40**:1—18.
2. **Paffenbarger R,** Blair S, Lee I. A history of physical activity, cardiovascular health and longevity: the scientific contributions of Jeremy N Morris, DSc, DPH, FRCP. *Int J Epidemiol* 2001;**30**:1184—92.
3. **Callaghan P.** Exercise: a neglected intervention in mental health care? *J Psychiatr Ment Health Nurs* 2004;**11**:476—83.
4. **Mead GE,** Morley W, Campbell P, *et al*. Exercise for depression. *Cochrane Database Syst Rev* 2009;(3):CD004366.
5. **Williams NH,** Hendry M, Frances B, *et al*. Effectiveness of exercise-referral schemes to promote physical activity in adults: systematic review. *Br J Gen Pract* 2007;**57**:979—86.
6. **Morgan O.** Approaches to increase physical activity: reviewing the evidence for exercise-referral schemes. *Public Health* 2005;**119**:361—70.
7. **National Institute for Health and Clinical Excellence.** *Public Health Collaborating Centre for Physical Activity. A Rapid Review of the Effectiveness of Exercise Referral Schemes to Promote Physical Activity in Adults*. London: National Institute for Health and Clinical Excellence, 2006.
8. **Gidlow C,** Johnston LH, Crone D, *et al*. Attendance of exercise referral schemes in the UK: a systematic review. *Health Educ J* 2005;**64**:168—86.
9. **Cobiac LJ,** Vos T, Barendregt JJ. Cost-effectiveness of interventions to promote physical activity: a modelling study. *PLoS Med* 2009;**6**:e1000110.
10. **Garrett S,** Raina E, Rose S, *et al*. Are physical activity interventions in primary care and the community cost-effective? A systematic review of the evidence. *Br J Gen Prac* 2011;**61**:125—33.
11. **Sowden SL,** Raine R. Running along parallel lines: how political reality impedes the evaluation of public health interventions. A case study of exercise referral schemes in England. *J Epidemiol Community Health* 2008;**62**:835—41.
12. **Murphy S,** Raisanen L, Moore G, *et al*. A pragmatic randomised controlled trial of the Welsh National Exercise Referral Scheme: protocol for trial and integrated economic and process evaluation *BMC Public Health* 2010;**10**:352.
13. **Department of Health.** *The General Practice Physical Activity Questionnaire*. London: Department of Health, 2006.
14. **Cock D,** Adams IC, Ibbetson AB, *et al*. REFERQUAL: a pilot study of a new service quality assessment instrument in the GP exercise referral scheme setting. *BMC Health Serv Res* 2006;**6**:61.
15. **Sykes K,** Roberts A. The Chester step test-a simple yet effective tool for the prediction of aerobic capacity. *Physiotherapy* 2004;**90**:183—8.
16. **Moore GF,** Moore L, Murphy S. Integration of motivational interviewing into practice in the national exercise referral scheme in Wales: a mixed methods study. *Behav Cogn Psychother* 2012;**40**:313—30.
17. **Blair SN,** Haskell WL, Paffenbarger RS, *et al*. Assessment of habitual physical activity by a seven-day recall in a community survey and controlled experiments. *Am J Epidemiol* 1985;**122**:794—804.
18. **Sarkin J,** Campbell J, Gross L, *et al*. Project GRAD seven day physical activity recall interviewer's manual. *Med Sci Sports Exerc* 1997;**29**(S6):S91—102.
19. **Knapp M,** Beecham J. Costing mental health services: the client service receipt inventory. *Psychol Med* 1990;**20**:893—908.
20. **Rabin R,** De Charro F. EQ-5D: a measure of health status from the EuroQol Group. *Ann Med* 2001;**33**:337—43.
21. **Morris S,** Devlin N, Parkin D. *Economic Analysis in Health Care*. UK: Wiley, 2007.
22. **Baecke JA,** Burema HJ, Frijters JE. A short questionnaire for the measurement of habitual physical activity in epidemiological studies. *Am J Clin Nutr* 1982;**36**:936—42.

23. **Snaith RP,** Zigmond AS. *HADS: Hospital Anxiety and Depression Scale*. 1994. Windsor: NFER Nelson, 1994.
24. **Welsh Assembly government.** *Welsh Index of Multiple deprivation*. Cardiff: Welsh Assembly Government, 2008. http://wales.gov.uk/docs/statistics/2009/090930wimd08laintroe.pdf (accessed 1 Feb 2010).
25. **NICE.** *Guide to Methods of Technology Appraisals*. 2008. http://www.nice.org.uk/aboutnice/howwework/devnicetech/technologyappraisalprocessguides/guidetothemethodsoftechnologyappraisal.jsp?domedia=1&mid=B52851A3-19B9-E0B5-D48284D172BD8459 (accessed 10 Aug 2011).
26. **Curtis L.** *Unit Costs of Health and Social Care*. UK: PSRRU, 2009. http://www.pssru.ac.uk/uc/uc2008contents.htm (accessed 9 Feb 2010).
27. **Department of Health.** *NHS reference Costs 2007-08*. 2009. http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_098945 (accessed 9 Feb 2010).
28. **NHS Information Centre.** *Prescribing cost analysis 2008 England*. 2009. http://www.ic.nhs.uk/statistics-and-data-collections/primary-care/prescriptions/prescription-cost-analysis-2008 (accessed 9 Feb 2010).
29. **Glick HA,** Doshi JA, Sonnad SS, *et al*. *Economic Evaluation in Clinical Trials (Handbooks for Health Economic Evaluation)*. 1st edn. Oxford: Oxford University Press, 2007.
30. **Sculpher M,** Gafni A. Recognizing diversity in public preferences: the use of preference sub-groups in cost-effectiveness analysis. *Health Econ* 2001;**10**:317—24.
31. **Pavey TG,** Taylor AH, Fox KR, *et al*. Effect of exercise referral schemes in primary care on physical activity and improving health outcomes: systematic review and meta-analysis. *BMJ* 2011;**343**:d6462.
32. **Pavey TG,** Anokye N, Taylor AH, *et al*. The clinical effectiveness and cost-effectiveness of exercise referral schemes: a systematic review and economic evaluation. *Health Technol Assess* 2011;**15**:i—xii, 1—254.
33. **Anokye NK,** Trueman P, Green C, *et al*. The cost-effectiveness of exercise referral schemes. *BMC Public Health* 2011;**11**:954.
34. **Moore GF,** Moore L, Murphy S. Facilitating adherence to physical activity: exercise professionals' experiences of the National Exercise Referral Scheme in Wales. *BMC Public Health* 2011;**11**:935.
35. **Issac AJ,** Critchley JA, Tai SS, *et al*. Exercise evaluation randomised trial (EXERT): a randomised trial comparing GP referral for leisure centre-based exercise, community-based walking and advice only. *Health Technol Assess* 2007;**11**:1—185.
36. **Green LW,** Glasgow RE. Evaluating the relevance, generalization, and applicability of research issues in external validation and translation methodology. *Eval Health Prof* 2006;**29**:126—53.
37. **BHF National Centre.** *A Toolkit for the design, implementation & Evaluationof exercise referral Schemes:Section 2*. Loughborough University, 2010. http://www.bhfactive.org.uk/sites/Exercise-Referral-Toolkit/downloads/s1-background-technical-report.pdf
38. **Macintyre S.** Good intentions and received wisdom are not good enough: the need for controlled trials in public health. *J Epidemiol Community Health* 2011;**65**:564—7.