OPEN ACCESS

University of BRISTOL

Donovan, J. L., Young, G. J., Walsh, E. I., Metcalfe, C., Lane, J. A., Martin, R. M., ... the ProtecT Study Group (2018). A prospective cohort and extended comprehensive-cohort design provided insights about the generalizability of a pragmatic trial: the ProtecT prostate cancer trial. *Journal of Clinical Epidemiology*, *96*, 35-46. https://doi.org/10.1016/j.jclinepi.2017.12.019

Publisher's PDF, also known as Version of record

License (if available):
CC BY-NC-ND

Link to published version (if available):
10.1016/j.jclinepi.2017.12.019

Link to publication record in Explore Bristol Research
PDF-document

ORIGINAL ARTICLE

A prospective cohort and extended comprehensive-cohort design provided insights about the generalizability of a pragmatic trial: the ProtecT prostate cancer trial

Jenny L. Donovan[a,b,*], Grace J. Young[a], Eleanor I. Walsh[a], Chris Metcalfe[a,c], J. Athene Lane[a,c], Richard M. Martin[a,d], Marta K. Tazewell[a], Michael Davis[a], Tim J. Peters[e], Emma L. Turner[a], Nicola Mills[a], Hanan Khazragui[a], Tarnjit K. Khera[a], David E. Neal[f], Freddie C. Hamdy[f], ProtecT Study Group

[a]School of Social and Community Medicine, University of Bristol, Bristol, UK
[b]NIHR Collaboration for Leadership in Applied Health Research and Care West, Hosted by University Hospitals Bristol NHS Foundation Trust, Bristol, UK
[c]Bristol Randomised Trials Collaboration, University of Bristol, Bristol, UK
[d]University Hospitals Bristol NHS Foundation Trust, National Institute for Health Research, Bristol Nutrition Biomedical Research Unit, Bristol, UK
[e]School of Clinical Sciences, University of Bristol, Bristol, UK
[f]Nuffield Department of Surgical Sciences, John Radcliffe Hospital, University of Oxford, Oxford, UK

Accepted 11 December 2017; Published online 27 December 2017

Abstract

Objectives: Randomized controlled trials (RCTs) deliver robust internally valid evidence but generalizability is often neglected. Design features built into the Prostate testing for cancer and Treatment (ProtecT) RCT of treatments for localized prostate cancer (PCa) provided insights into its generalizability.

Study Design and Setting: Population-based cluster randomization created a prospective study of prostate-specific antigen (PSA) testing and a comprehensive-cohort study including groups choosing treatment or excluded from the RCT, as well as those randomized. Baseline information assessed selection and response during RCT conduct.

Results: The prospective study (82,430 PSA-tested men) represented healthy men likely to respond to a screening invitation. The extended comprehensive cohort comprised 1,643 randomized, 997 choosing treatment, and 557 excluded with advanced cancer/comorbidities. Men choosing treatment were very similar to randomized men except for having more professional/managerial occupations. Excluded men were similar to the randomized socio-demographically but different clinically, representing less healthy men with more advanced PCa.

Conclusion: The design features of the ProtecT RCT provided data to assess the representativeness of the prospective cohort and generalizability of the findings of the RCT. Greater attention to collecting data at the design stage of pragmatic trials would better support later judgments by clinicians/policy-makers about the generalizability of RCT findings in clinical practice. © 2018 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Keywords: Randomized; Clinical trial; Generalizability; External validity; Prostate cancer; Comprehensive cohort

What is new?

Key findings

- Decisions taken when designing the ProtecT prostate cancer (PCa) treatment trial and linked Cluster randomized trial of PSA testing for prostate cancer screening trial enabled the collection of data to assess the representativeness of the prospective study of prostate-specific antigen testing and generalizability of the findings of the ProtecT RCT.

- Adding the extended comprehensive-cohort study comprising all men diagnosed with PCa, including those who chose a treatment or were ineligible for the RCT with advanced cancer or comorbidities as well as those randomized in ProtecT, allowed the assessment of the generalizability of the trial's findings to patients diagnosed with PCa in routine care.

What this adds to what was known?

- Aspects of the generalizability of pragmatic RCTs can be evaluated through initiatives at the design phase, such as assessing factors associated with participation at various stages through a preceding prospective study and/or collecting data from those choosing treatments or excluded from the trial according to eligibility criteria, although these decisions will have time and resource implications.

- Including an innovatively extended comprehensive-cohort study of all men diagnosed with a condition like PCa can enable assessment of important similarities and differences between the randomized group and those who choose a treatment in standard practice or with aspects of advanced cancer or comorbidities that preclude trial participation, providing insights into the applicability of the RCT to patients in routine practice.

What is the implication and what should change now?

- Greater attention should be devoted at the design stage of pragmatic RCTs to ensure that appropriate data are collected to support later judgments by clinicians and policy-makers about the generalizability of the findings of an RCT to patients in routine clinical practice.

1. Introduction

1.1. Pragmatic RCTs and generalizability

Randomized controlled trials (RCTs) offer the most rigorous way to evaluate the effectiveness of treatments, but there are often concerns about the generalizability of findings [1–3]. A real or perceived lack of relevance to patients in routine care contributes to the slow or limited uptake of RCT evidence into practice [4,5]. RCTs remain the primary design for evaluation because random allocation of participants to treatment groups helps ensure against selection bias. Whether and to what degree the findings of an RCT can then be generalized to patients in similar or different settings or with different but related disease characteristics requires judgments, including reflection on the evidence from a new study in relation to prior knowledge, statistical reasoning, biological plausibility, and interpretations of the impact of eligibility criteria of the RCT in the context of contemporary clinical practice [3].

Decisions at the design stage of an RCT can facilitate or inhibit later judgments about the generalizability and clinical relevance of the findings. The PRagmatic Explanatory Continuum Indicator Summary-2 (PRECIS-2) tool was developed to support trialists in making decisions to position an RCT along the continuum between explanatory efficacy approaches in ideal circumstances and pragmatic designs evaluating effectiveness within "real-world" naturalistic settings [6]. There is consensus about the value of pragmatic designs in informing clinical decision-making, but there has been considerable recent debate in this journal about how best to design such RCTs and the role of the PRECIS-2 tool [7–11]. In addition, a recent series has sought to provide theoretical and practical guidance to promote operational feasibility in pragmatic RCTs [12,13].

Pragmatic RCTs need to closely resemble the population and clinical practice they aim to influence to ensure they retain the advantages of randomization while adding the ability to produce findings that are generalizable. However, during the implementation of such RCTs, unanticipated challenges often arise in relation to recruitment, setting, equipoise, or other aspects of RCT conduct, or changes in clinical practice [10]. Judgments about the generalizability of an RCT require knowledge about local health and care conditions, and trialists may not be best placed to do this [7]. To ensure that evidence-based judgments can be made, trialists also need to ensure that their design allows the collection of data that will later facilitate the assessment of the RCT findings' generalizability and clinical relevance. We were able to collect such data in linked RCTs in the area of prostate cancer (PCa).

1.2. Evaluating screening and treatment for prostate cancer

PCa is a major cause of death for older men, and although the prostate-specific antigen (PSA) blood test provides the opportunity to identify the disease at a stage when it could be cured, screening detects many tumors that will not become clinically important and so receive unnecessary radical/curative treatments that cause damaging side-effects. Previously published RCTs have focused either on screening or treatment and have contributed valuable knowledge but not provided consistent findings [14—17]. In the mid/late 1990s, we designed two interlinked pragmatic RCTs aiming to inform policy for PCa screening and treatment practice:

(a) Cluster randomized trial of PSA testing for prostate cancer (CAP): The Cancer Research UK/UK Department of Health, CAP employed cluster randomization of general practices in a Zelen design to create an intervention arm comprising a prospective study of

men undergoing PSA testing and a control arm of usual National Health Service (NHS) care without organized PSA testing (Fig. 1; baseline details [18]).

(b) ProtecT: The National Institute for Health Research, ProtecT (Prostate testing for cancer and Treatment) RCT evaluated the effectiveness and cost-effectiveness of the three major standard treatments for clinically localized PCa diagnosed during the prospective study of PSA testing: radical surgery, radical external-beam radiotherapy, and active monitoring (Fig. 1; baseline details [19]).

These RCTs were intended to be pragmatic in design, and knowing that the primary outcomes would not be published until a median of 10 years' follow-up, decisions were taken at the design stage to provide data to facilitate the later evaluation of the generalizability and clinical relevance of the findings.

Initially, a feasibility study was undertaken to investigate whether it was possible to recruit men from the community to
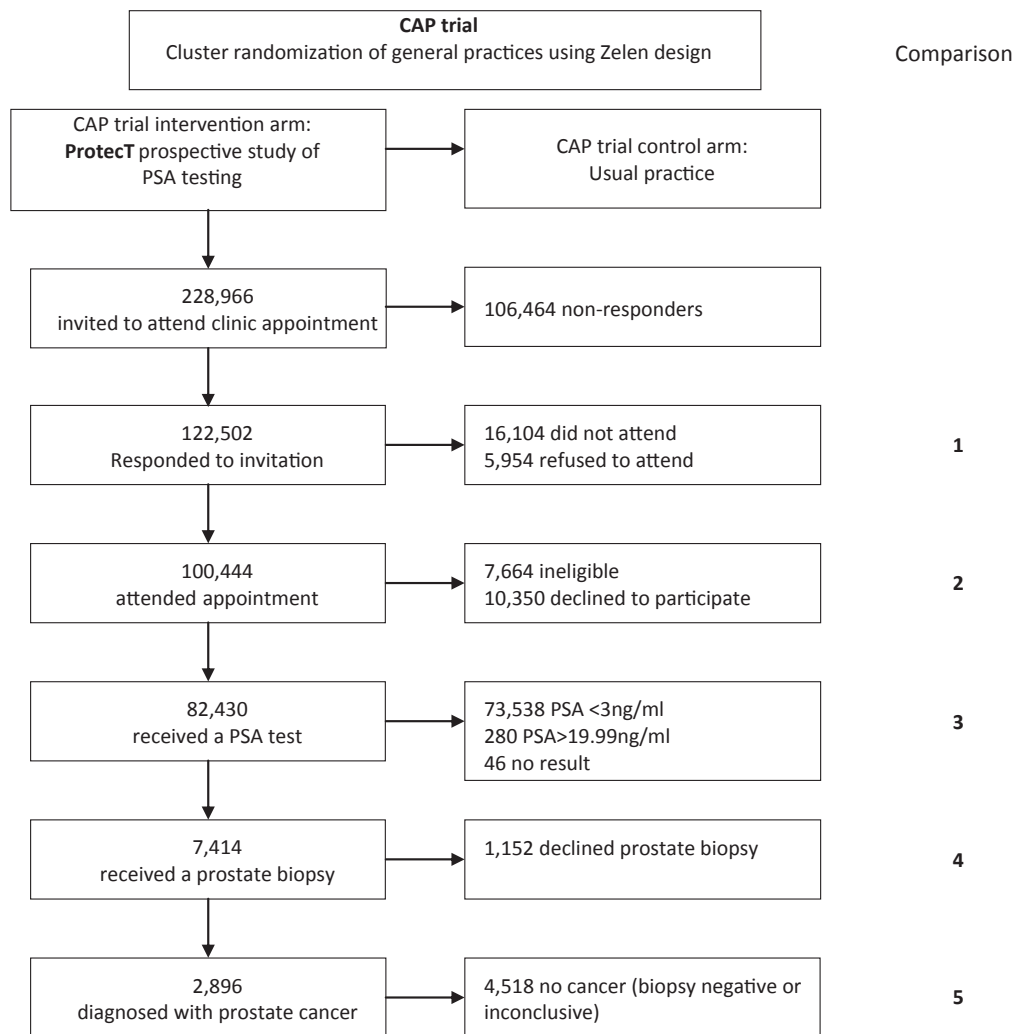

Fig. 1. CAP trial framework and comparison points in the ProtecT prospective study of PSA testing and diagnosis. PSA, prostate-specific antigen; CAP, Cluster randomized trial of PSA testing for prostate cancer.

have a PSA test and then randomize those diagnosed with clinically localized PCa into a treatment trial. When this feasibility was assured [20], the CAP RCT was initiated to evaluate screening. The population-based cluster design of CAP created an intervention arm comprising a prospective study of PSA testing within which the ProtecT RCT of treatments was embedded (Fig. 1). As recruitment to ProtecT was anticipated to be particularly challenging because of randomization between surgery, radiotherapy or no immediate treatment (active monitoring), an integrated recruitment study was undertaken [21] and a comprehensive-cohort study as in [22] to follow-up men who declined randomization and chose a treatment alongside those who agreed to be randomized. The comprehensive cohort was then extended to include all men diagnosed with PCa during the prospective study but excluded from the treatment trial because of advanced PCa or comorbidity—many of these men would have received one of the study treatments in usual practice (although they would not be eligible for all three as in the RCT).

The collection of individual participant socio-demographic, symptomatic, and clinical data at baseline in the prospective study of PSA testing enabled the investigation of the representativeness of the study population and selection factors at each stage of response and clinical eligibility. These data provide information to enable judgments about the generalizability and clinical relevance of the findings of the recently published ProtecT primary outcomes [23,24]. This article presents the data generated by the CAP/ProtecT design features, their limitations, and the insights that they can provide; with brief consideration of the value and practicality of such design features in pragmatic RCTs more generally to inform assessments of generalizability.

2. Methods

2.1. The CAP randomized controlled trial

Cluster randomization of more than 900 primary-care centers in the UK created an intervention arm of men aged 50–69 years invited to PSA testing and a control arm of usual NHS care without organized PSA testing, followed-up using routinely collected mortality data [18] (Fig. 1). There was no evidence of differences between primary-care centers agreeing or declining to participate in CAP or between men in the intervention and control practices [25]. The CAP intervention arm provided a population-based framework for the recruitment of men into the prospective study of PSA testing and ProtecT treatment RCT.

2.2. ProtecT prospective study of PSA tesing

Men aged 50–69 years registered in primary-care centers were sent one invitation to attend an appointment to discuss PSA testing and the ProtecT RCT. Data available to compare responders and nonresponders to the

appointment and PSA testing were restricted to date of birth and postcode. While men who responded to the invitation were similar to nonresponders, except for being slightly less deprived [26], data to evaluate more detailed characteristics of nonresponders were not available. Men attending an appointment who consented to a blood test for PSA had socio-demographic and clinical history information collected and completed a brief study questionnaire, with a more detailed questionnaire requested from men later undergoing prostate biopsies [27]. Comparisons were made between those participating or not at each stage through exclusion or choice, to provide insights about representativeness and generalizability.

2.3. ProtecT RCT recruitment and comprehensive-cohort study

Men diagnosed with clinically localized PCa through the PSA-testing study and meeting the inclusion/exclusion criteria were eligible for recruitment to the ProtecT RCT. They attended an appointment with an urologist for the diagnosis and basic information and received detailed information about treatment options and the ProtecT RCT from a research nurse. Men were asked if they would consent to random allocation of treatment. If they declined randomization, they chose a treatment and were included in the comprehensive cohort (Fig. 2), followed-up identically to the randomized.

2.4. ProtecT extended comprehensive-cohort study

Men diagnosed with PCa but excluded from the RCT because they had advanced cancer or were not eligible for all three treatments were included in the extended comprehensive cohort (Fig. 2).

2.5. Statistical analysis

Socio-demographic information and data from the participant questionnaires were used to compare the baseline characteristics of the men at different stages of response and eligibility in the prospective PSA-testing study to assess response and selection (comparisons 1–5 Fig. 1) and to compare the randomized group with those choosing their treatment, diagnosed with advanced cancer or excluded from the RCT for other reasons (comparisons 6–8, Fig. 2).

All statistical analyses were completed using STATA, version 14.1. For continuous socio-demographic variables, medians and interquartile ranges were reported with Mann–Whitney tests to analyze differences between groups. Questionnaire data were presented with means and standard deviations. Between-group comparisons were carried out to investigate whether baseline characteristics differed between those proceeding through PSA testing and PCa diagnosis or excluded or choosing not to participate in the prospective cohort; and between the randomized and other groups in the extended comprehensive cohort. With ceiling effects
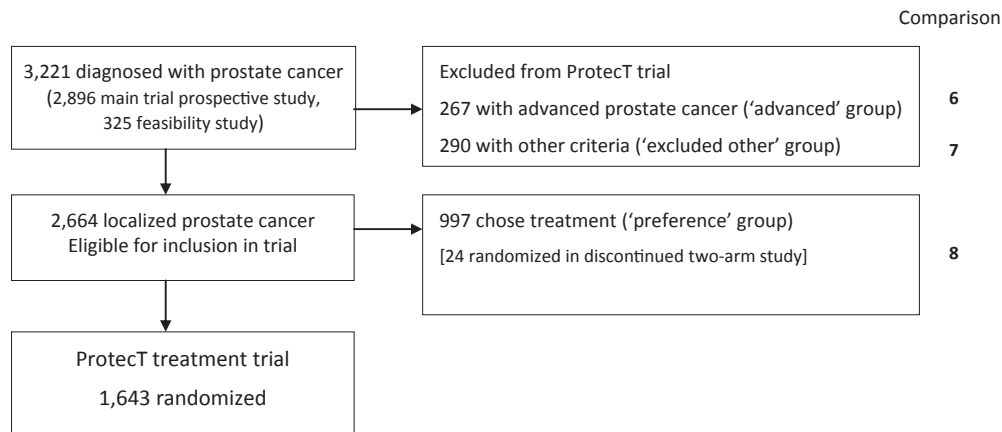
Comparison

Fig. 2. ProtecT RCT extended comprehensive-cohort study and comparison points. ProtecT, prostate testing for cancer and treatment; RCT, randomized controlled trial.

evident at baseline, each continuous comparison was tested using a nonparametric Mann−Whitney test. For binary variables, such as previous PSA test, groups were compared using logistic regression. Ordered categorical variables such as occupation and cancer staging were analyzed using ordinal logistic regression with the most-desirable/least-worse category as the base comparator. Adjustment for age and center in the logistic and ordinal logistic models did not influence overall conclusions. Given the sample size and large number of tests, greater attention was given to descriptive statistics rather than *P*-values: for continuous variables, we considered with interest but caution differences ≥ 0.5 standard deviations; similarly for categorical variables with differences that resulted in a risk ratio of ≤ 0.9.

3. Results

3.1. Prospective study of PSA testing and PCa diagnosis

In total, 122,502 men responded to the invitation for a PSA test, and 100,444 attended. Those who explicitly refused to attend (5,954) or did not attend after agreeing to do so (16,104) lived in more deprived areas than attenders (comparison 1, Fig. 1, Table Web 1 on the journal's website at www.elsevier.com). Men who attended but declined the PSA test (10,350) or were ineligible (7,665) also lived in more deprived areas than those who attended (comparison 2, Fig. 1). Eighty-two thousand four hundred thirty men attended and received a PSA test (36% of those invited) (Fig. 1, Tables 1 and 2 and Web 1 on the journal's website at www.elsevier.com).

At each of the stages of PSA testing, biopsy, and PCa diagnosis, the groups eligible for the RCT were very similar to those excluded in terms of socio-demographic characteristics (comparisons 3−5 Fig. 1). Expected clinical relationships were found, such as a positive relationship between PSA and age, and weak evidence for family history of the disease and diagnosis (Table 1). It was notable that those more likely to have a high PSA-test result or

diagnosis of PCa were less likely to have previously had a PSA test or urological treatment (Table 1).

3.2. ProtecT recruitment and comprehensive cohort

Overall, 3,221 men were diagnosed with PCa: 2,896 in the prospective study and 325 during the feasibility phase (Fig. 2). A total of 2,664 (83%) had clinically localized PCa (stage T1/T2) and were eligible for inclusion in the ProtecT RCT. An integrated recruitment study was undertaken to understand the issues underlying recruitment difficulties and provide improvements to study information and presentation. This increased the percentage consenting to randomization from 30% in the early stages to 62% at completion [21]. The comprehensive cohort comprised:

- ProtecT randomized cohort: 1,643 men (62%) who consented to randomization to the ProtecT RCT comparing active monitoring, surgery, and radiotherapy.
- ProtecT "treatment-choice" cohort: 997 men (38%) who declined randomization and chose their treatment (273 surgery, 133 radiotherapy, 529 active monitoring, and 62 other options not included in the RCT [brachytherapy and high-frequency ultrasound]).

The "treatment-choice" group was very similar to the randomized in relation to clinical and socio-demographic characteristics, except that those who chose treatment were more likely to be in managerial/professional occupations than the randomized (53% vs. 42%), less deprived (11% vs. 15%) (Table 3), and more likely to have had a previous PSA test (18% vs. 14%). The groups were almost identical in response to general health measures and symptom scores (Table 4).

3.3. Extended comprehensive-cohort study

The following were excluded from randomization:

- Two hundred sixty-seven (8%) with advanced cancer (stage T3 or higher).

Table 1. Baseline differences in socio-demographic and clinical factors for participants in the ProtecT prospective study of PSA testing and cancer diagnosis

			Had a PSA test[a] $n = 82,430$				
				Trial eligible PSA ($3 \leq PSA < 20$) $n = 8,566$			
						Biopsy $n = 7,414$	
			All eligible $n = 8,566$	Declined biopsy $n = 1,152$	Received biopsy $n = 7,414$	Negative $n = 4,518$	Positive $n = 2,896$
	PSA <3 $n = 73,538$	PSA ≥ 20 $n = 280$					
Age (n)	73,538	280	8,566	1,152	7,414	4,518	2,896
Median age (IQR)	58.0 (8.0)[b]	64.0 (7.0)	62.0 (8.0)[b]	62.0 (8.0)	62.0 (8.0)	61.0 (8.0)	62.0 (8.0)
P value	$P < 0.001$[c]	$P < 0.001$[d]		$P < 0.001$[e]		$P < 0.001$[f]	
Ethnicity							
White n (%)	71,948 (98)	265 (96)	8,377 (99)	1,113 (98)	7,264 (99)	4,425 (99)	2,839 (99)
Other n (%)	1,127 (2)	10 (4)	108 (1)	23 (2)	85 (1)	52 (1)	33 (1)
P value	$P = 0.055$[c]	$P = 0.001$[d]		$P = 0.017$[e]		$P = 0.961$[f]	
Marital status							
Married/living as married n (%)	61,507 (84)	226 (82)	7,091 (83)	941 (83)	6,150 (84)	3,730 (83)	2,420 (84)
Other (divorced) n (%)	11,641 (16)	49 (18)	1,410 (17)	198 (17)	1,212 (16)	755 (17)	457 (16)
P value	$P = 0.110$[c]	$P = 0.589$[d]		$P = 0.437$[e]		$P = 0.284$[f]	
Occupation present or last paid							
Managerial n (%)	9,948 (44)	106 (41)[b]	3,783 (46)[b]	499 (49)	3,284 (46)	2,024 (47)	1,260 (44)
Intermediate n (%)	3,886 (17)	44 (17)	1,351 (16)	155 (15)	1,196 (17)	734 (17)	462 (16)
Working n (%)	8,717 (39)	110 (42)[b]	3,067 (37)[b]	355 (35)	2,712 (38)	1,588 (37)	1,124 (39)
P value	$P = 0.005$[c]	$P = 0.073$[d]		$P = 0.036$[e]		$P = 0.018$[f]	
Cancer/treatment history							
Previous PSA test (%)	9,229 (13)[b]	17 (6)[b]	1,594 (19)[b]	279 (25)[b]	1,315 (18)[b]	892 (20)[b]	423 (15)[b]
P value	$P < 0.001$[c]	$P < 0.001$[d]		$P < 0.001$[e]		$P < 0.001$[f]	
Previous urinary/ prostate treatment (%)	5,980 (8)[b]	24 (9)[b]	1,069 (13)[b]	171 (15)[b]	898 (12)[b]	644 (14)[b]	254 (9)[b]
P value	$P < 0.001$[c]	$P = 0.058$[d]		$P = 0.007$[e]		$P < 0.001$[f]	
Family history of cancer (prostate only) (%)	3,748 (6)[b]	17 (7)	554 (7)[b]	70 (7)	484 (7)	264 (7)[b]	220 (9)[b]
P value	$P < 0.001$[c]	$P = 0.765$[d]		$P = 0.652$[e]		$P = 0.002$[f]	
Family history of cancer (all) (%)	36,541 (53)	144 (57)	4,445 (56)	569 (54)	3,876 (56)	2,320 (55)	1,556 (57)
P value	$P < 0.001$[c]	$P = 0.567$[d]		$P = 0.177$[e]		$P = 0.030$[f]	
Deprivation score (overall n)	73,027	279	8,488	1,143	7,345	4,484	2,861
Living in an area of deprivation n (%)	10,016 (14)	34 (12)	1,097 (13)	152 (13)	945 (13)	538 (12)[b]	407 (14)[b]
P value	$P = 0.044$[c]	$P = 0.718$[d]		$P = 0.685$[e]		$P = 0.005$[f]	
PSA level at baseline (n)	73,538	280	8,566	1,152	7,414	4,518	2,896
Median PSA level (IQR)	0.9 (0.9)[b]	32.1 (29.0)[b]	4.2 (2.5)[b]	3.9 (2.0)	4.3 (2.5)	4.1 (2.0)	4.8 (3.4)
P value				$P < 0.001$[e]		$P < 0.001$[f]	

Abbreviations: IQR, interquartile range; ProtecT, prostate testing for cancer and treatment; PSA, prostate-specific antigen; RR, risk ratio.
[a] Forty-six men had no result.
[b] Differences of interest when using the ≥ 0.5 SDS cutoff for continuous outcomes or the ≤ 0.9 RR cutoff for categorical outcomes.
[c] Comparison between PSA <3 and eligible PSA.
[d] Comparison between PSA ≥ 20 and eligible PSA.
[e] Comparison between those who had a biopsy and those who did not.
[f] Comparison between those with a negative biopsy result and those with a positive biopsy result.

• Two hundred ninety (9%) because they were considered unsuitable for the treatments for other reasons, mostly comorbidities.

As expected, the 267 "advanced cancer" group had much higher PSA levels, cancer stage (95% T3), and PCa grade (71% Gleason 7 or more) than those randomized

Table 2. Patient-reported general health and symptomatic measures: baseline differences for participants in the ProtecT prospective study of PSA testing

	PSA <3 n = 73,538	PSA ≥20 n = 280	All eligible n = 8,566	Declined biopsy n = 1,152	Received biopsy n = 7,414	Negative n = 4,518	Positive n = 2,896
SF-12 (minimum n)	60,146	225	6,925	902	6,023	3,710	2,313
Mean physical score (SD)	49.6 (9.0)	49.5 (9.2)	49.4 (9.0)	48.6 (9.6)	49.5 (8.9)	49.7 (8.8)	49.3 (9.0)
P value	P = 0.029[b]	P = 0.764[c]		P = 0.028[d]		P = 0.070[e]	
Mean mental score (SD)	53.2 (8.4)	54.3 (7.8)	53.8 (7.8)	53.9 (8.1)	53.8 (7.8)	53.8 (7.8)	53.8 (7.8)
P value	P < 0.001[b]	P = 0.205[c]		P = 0.219[d]		P = 0.715[e]	
HADS (minimum n)	60,917	238	7,241	960	6,281	3,818	2,459
Anxiety case (≥8) n (%)	13,882 (23)[f]	46 (19)	1,413 (20)[f]	168 (17)[f]	1,245 (20)[f]	749 (20)	496 (20)
P value	P < 0.001[b]	P = 0.920[c]		P = 0.086[d]		P = 0.613[e]	
Depression case (≥8) n (%)	4,905 (8)[f]	13 (5)[f]	508 (7)[f]	80 (8)[f]	428 (7)[f]	244 (6)[f]	184 (7)[f]
P value	P = 0.002[c]	P = 0.356[c]		P = 0.087[d]		P = 0.092[e]	
Mean anxiety score (SD)	5.2 (3.5)	4.5 (3.5)	4.9 (3.4)	4.7 (3.5)	4.9 (3.3)	4.9 (3.3)	5.0 (3.4)
P value	P < 0.001[c]	P = 0.040[c]		P = 0.026[d]		P = 0.420[e]	
Mean depression score (SD)	3.1 (2.8)	3.0 (2.7)	3.0 (2.7)	3.0 (3.0)	3.0 (2.6)	2.9 (2.6)	3.1 (2.7)
P value	P = 0.012[c]	P = 0.912[c]		P = 0.299[d]		P = 0.076[e]	
EQ5D (n)	66,332	257	7,744	1,020	6,724	4,076	2,648
Mean EQ5D score	0.9 (0.2)	0.9 (0.2)	0.9 (0.2)	0.9 (0.2)	0.9 (0.2)	0.9 (0.2)	0.9 (0.2)
P value	P = 0.040[c]	P = 0.854[c]		P = 0.499[d]		P = 0.003[e]	
ICSmaleSF − symptoms (minimum n)	67,084	258	7,883	1,028	6,851	4,173	2,678
Delay before urinating n (%)	29,642 (44)[f]	136 (52)	4,340 (55)[f]	522 (50)	3,818 (55)	2,435 (58)[f]	1,383 (51)[f]
P value	P < 0.001[c]	P = 0.340[c]		P = 0.003[d]		P < 0.001[e]	
Rush to the toilet n (%)	30,586 (45)[f]	135 (51)[f]	4,549 (57)[f]	554 (53)	3,995 (58)	2,546 (60)[f]	1,449 (54)[f]
P value	P < 0.001[b]	P = 0.051[c]		P = 0.003[d]		P < 0.001[e]	
Leak before reaching the toilet n (%)	13,932 (21)[f]	82 (31)	2,246 (28)[f]	279 (27)	1,967 (29)	1,298 (31)[f]	669 (25)[f]
P value	P < 0.001[b]	P = 0.323[c]		P = 0.225[d]		P < 0.001[e]	
Frequency (≤3 h per void) n (%)	43,699 (65)	171 (66)	5,496 (70)	686 (66)	4,810 (70)	2,994 (72)	1,816 (68)
P value	P < 0.001[b]	P = 0.238[c]		P = 0.015[d]		P = 0.001[e]	
Nocturia n (%)	45,310 (67)[f]	183 (70)	5,915 (75)[f]	756 (74)	5,159 (75)	3,167 (76)	1,992 (74)
P value	P < 0.001[b]	P = 0.059[c]		P = 0.245[d]		P = 0.196[e]	
Do urinary symptoms interfere with life? n (%)	13,466 (20)[f]	60 (23)[f]	2,272 (29)[f]	270 (26)[f]	2,002 (29)[f]	1,355 (32)[f]	647 (24)[f]
P value	P < 0.001[b]	P = 0.037[c]		P = 0.032[d]		P < 0.001[e]	

Abbreviations: IQR, interquartile range; ProtecT, prostate testing for cancer and treatment; PSA, prostate-specific antigen; SD, standard deviation; RR, risk ratio; SF-12, Short-form 12-item questionnaire; EQ5D, Euroqol 5 dimension scale.

[a] Forty-six men had no result.
[b] Comparison between PSA <3 and eligible PSA.
[c] Comparison between PSA ≥20 and eligible PSA.
[d] Comparison between those who had a biopsy and those who did not.
[e] Comparison between those with a negative biopsy result and those with a positive biopsy result.
[f] Differences of interest when using the ≥0.5 SDS cutoff for continuous outcomes or the ≤0.9 RR cutoff for categorical outcomes.

(Table 3). They were very similar to the randomized group in terms of socio-demographic characteristics and health and symptom scores, although much less likely to have had a previous PSA test (7% vs. 14%) (Table 3). The 290 ''excluded other'' group had higher grade (37% vs. 23% Gleason 7 or higher) and stage (38% vs. 24% T2) PCa than those randomized, although not as high as the ''advanced cancer'' men. They were similar to the randomized group in terms of socio-demographic characteristics and clinical history, although more likely to have had previous urological treatment (Table 3) and as expected, slightly poorer health status with more depression and some worse urinary symptoms (Table 4) (Fig. 2, Tables 3 and 4).

4. Discussion

The major aim of pragmatic RCTs is to produce findings that are clinically relevant and generalizable beyond the specific participants. The ProtecT RCT was designed to evaluate the effectiveness of treatments for clinically localized PCa and was embedded in the intervention arm of the CAP RCT evaluating population screening. The ProtecT RCT was designed in the late 1990s, more than 15 years before the results were published [23,24], aiming to be pragmatic and with design features that provided opportunities to collect data that can now be used to produce insights into the representativeness of the PSA-tested cohort

Table 3. Baseline differences in socio-demographic and clinical factors for those participating in the ProtecT RCT extended comprehensive-cohort study

	Randomized $n = 1,643$	Advanced $n = 267$	Excluded (other) $n = 290$	Choosing treatment $n = 997$
Age (n)	1,643	267	290	997
Median age (IQR)	62.0 (9.0)	63.0 (8.0)	63.0 (8.0)	62.0 (7.0)
P value		$P = 0.001^a$	$P < 0.001^b$	$P = 0.604^c$
Ethnicity				
White n (%)	1,606 (99)	259 (99)	283 (99)	984 (99)
Other n (%)	22 (1)	3 (1)	3 (1)	9 (1)
P value		$P = 0.786^a$	$P = 0.679^b$	$P = 0.310^c$
Marital status				
Married/living as married n (%)	1,375 (84)	231 (88)	232 (81)	841 (85)
Other (e.g. divorced) n (%)	257 (16)	31 (12)	56 (19)	151 (15)
P value		$P = 0.103^a$	$P = 0.118^b$	$P = 0.719^c$
Occupation present or last paid				
Managerial n (%)	684 (42)^d	107 (42)	121 (44)	516 (53)^d
Intermediate n (%)	259 (16)^d	45 (18)	46 (17)	157 (16)^d
Working n (%)	678 (42)^d	104 (41)	111 (40)	307 (31)^d
P value		$P = 0.898^a$	$P = 0.589^b$	$P < 0.001^c$
Cancer/treatment history				
Previous PSA test (%)	227 (14)^d	19 (7)^d	37 (13)	175 (18)^d
P value		$P = 0.004^a$	$P = 0.720^b$	$P = 0.012^c$
Previous urinary/prostate treatment (%)	142 (9)^d	26 (10)^d	36 (13)^d	82 (8)^d
P value		$P = 0.510^a$	$P = 0.040^b$	$P = 0.694^c$
Family history of cancer (prostate only) (%)	119 (8)^d	19 (9)^d	21 (8)	83 (9)^d
P value		$P = 0.829^a$	$P = 0.926^b$	$P = 0.271^c$
Family history of cancer (all) (%)	897 (58)	131 (55)	144 (54)	543 (58)
P value		$P = 0.388^a$	$P = 0.235^b$	$P = 0.885^c$
Deprivation score (n)	1,624	262	285	977
Living in an area of deprivation n (%)	239 (15)^d	39 (15)	44 (15)	111 (11)^d
P value		$P = 0.943^a$	$P = 0.752^b$	$P = 0.015^c$
PSA level (minimum n)	1,631	167	198	876
Median baseline PSA level (IQR)	4.6 (3.1)^d	8.5 (8.0)^d	5.2 (4.8)	4.8 (3.1)
P value		$P < 0.001^a$	$P = 0.002^b$	$P = 0.455^c$
Median biopsy PSA level (IQR)	4.8 (3.4)^d	9.1 (8.8)^d	5.9 (4.8)	4.8 (3.3)
P value		$P < 0.001^a$	$P < 0.001^b$	$P = 0.714^c$
Gleason score				
6 (%)	1,266 (77)^d	75 (28)^d	181 (63)^d	755 (76)
7 (%)	339 (21)^d	140 (52)^d	86 (30)^d	218 (22)
8−10 (%)	37 (2)^d	52 (19)^d	19 (7)^d	24 (2)
P value		$P < 0.001^a$	$P < 0.001^b$	$P = 0.419^c$
Cancer staging				
T1~ (%)	1,249 (76)^d	5 (2)^d	172 (61)^d	758 (76)
T2 (%)	394 (24)^d	3 (1)^d	106 (38)^d	239 (24)
T3^ (%)	0 (0)^d	250 (95)^d	3 (1)^d	0 (0)
T4+ (%)	0 (0)^d	4 (2)^d	0 (0)^d	0 (0)
P value		$P < 0.001^a$	$P < 0.001^b$	$P = 0.996^c$
Risk categorization^e				
Low (%)	951 (58)	0 (0)	110 (38)	559 (56)
Intermediate (%)	654 (40)	215 (81)	159 (55)	414 (42)
High (%)	37 (2)	52 (19)	19 (7)	24 (2)
P value		$P < 0.001^a$	$P < 0.001^b$	$P = 0.350^c$

Abbreviations: ProtecT, prostate testing for cancer and treatment; PSA, prostate-specific antigen; RCT, randomized controlled trial; RR, risk ratio.

~Includes 18 ''T1/T2'' in the advanced or excluded other group; ^ includes two ''T3/T4'' in the advanced or excluded other group.

[a] Comparison between randomized and advanced men.

[b] Comparison between randomized and excluded men.

[c] Comparison between randomized and preference men.

[d] Differences of interest when using the ≥ 0.5 SDS cutoff for continuous outcomes or the ≤ 0.9 RR cutoff for categorical outcomes.

[e] Defined as ''low'' if T1 and G ≤ 6 and PSA < 10, ''high'' if G ≥ 8, ''intermediate'' for all other combinations of stage, grade, and PSA.

in relation to the general population of men aged 50−69 years and the generalizability of the ProtecT treatment RCT findings to patients diagnosed with PCa in clinical practice.

The CAP cluster randomization of primary-care practices created comparable intervention and control arms, and so men invited to the prospective study of PSA testing were representative of the population of men aged

Table 4. Patient-reported general health and symptomatic measures: baseline differences for those participating in the ProtecT RCT extended comprehensive-cohort study

	Randomized $n = 1,643$	Advanced $n = 267$	Excluded (other) $n = 290$	Choosing treatment $n = 997$
SF-12 (minimum *n*)	1,260	172	190	778
Mean physical score (SD)	51.2 (7.9)	50.0 (8.7)	47.6 (10.8)	51.3 (7.9)
P value		$P = 0.142^a$	$P < 0.001^b$	$P = 0.464^c$
Mean mental score (SD)	53.9 (7.5)	53.4 (8.4)	53.4 (8.9)	53.5 (8.2)
P value		$P = 0.982^a$	$P = 0.875^b$	$P = 0.974^c$
HADS (minimum *n*)	1,399	201	228	853
Anxiety case ($\geq$8) *n* (%)	278 (20)	47 (23)	51 (22)	180 (21)
P value		$P = 0.317^a$	$P = 0.373^b$	$P = 0.504^c$
Depression case ($\geq$8) *n* (%)	80 (6)^d	12 (6)	21 (9)^d	44 (5)^d
P value		$P = 0.886^a$	$P = 0.045^b$	$P = 0.572^c$
Mean anxiety score (SD)	4.9 (3.5)	5.2 (3.8)	5.1 (3.8)	4.9 (3.5)
P value		$P = 0.328^a$	$P = 0.531^b$	$P = 0.933^c$
Mean depression score (SD)	2.5 (2.5)	2.6 (2.8)	3.2 (3.1)	2.5 (2.6)
P value		$P = 0.844^a$	$P = 0.001^b$	$P = 0.886^c$
EQ5D (*n*)	1,413	206	224	854
Mean EQ5D score	0.9 (0.2)^d	0.9 (0.2)	0.8 (0.2)^d	0.9 (0.2)
P value		$P = 0.022^a$	$P < 0.001^b$	$P = 0.260^c$
ICSmaleSF – symptoms (minimum *n*)	1,410	208	230	856
Delay before urinating *n* (%)	725 (51)	101 (48)	130 (56)	422 (49)
P value		$P = 0.388^a$	$P = 0.160^b$	$P = 0.387^c$
Rush to the toilet *n* (%)	844 (59)	137 (65)	149 (64)	502 (58)
P value		$P = 0.098^a$	$P = 0.199^b$	$P = 0.596^c$
Leak before reaching the toilet *n* (%)	407 (29)^d	72 (34)^d	81 (35)^d	206 (24)^d
P value		$P = 0.081^a$	$P = 0.060^b$	$P = 0.015^c$
Frequency ($\leq$2 h per void) *n* (%)	460 (33)	66 (32)	72 (31)	245 (30)
P value		$P = 0.797^a$	$P = 0.632^b$	$P = 0.143^c$
Nocturia ($>$1 per night) *n* (%)	312 (22)^d	63 (30)^d	65 (28)^d	166 (19)^d
P value		$P = 0.010^a$	$P = 0.034^b$	$P = 0.147^c$
Do urinary symptoms interfere with life? *n* (%)	367 (26)	58 (28)	66 (28)	211 (24)
P value		$P = 0.531^a$	$P = 0.401^b$	$P = 0.489^c$
ICSmaleSF scales (minimum *n*)	1,413	207	231	854
Mean ICSmaleVS (voiding scale)	3.3 (3.0)	3.2 (3.1)	3.8 (3.6)	3.3 (3.3)
P value		$P = 0.549^a$	$P = 0.157^b$	$P = 0.310^c$
Mean ICSmaleIS (incontinence scale)	1.8 (1.9)	1.9 (1.8)	2.2 (2.2)	1.6 (1.7)
P value		$P = 0.668^a$	$P = 0.074^b$	$P = 0.087^c$
ICIQ (*n*)	1,244	174	202	757
Mean ICIQ (SD)	1.3 (2.3)	1.3 (2.2)	1.6 (2.6)	1.0 (2.1)
P value		$P = 0.817^a$	$P = 0.093^b$	$P = 0.017^c$
ICIQ QoL impact: none (%)	1,174 (93)	168 (94)	183 (89)	725 (95)
ICIQ QoL impact: moderate (%)	81 (6)	11 (6)	22 (11)	39 (5)
ICIQ QoL impact: high (%)	4 ($<$1)	0 (0)	0 (0)	1 ($<$1)
P value		$P = 0.754^a$	$P = 0.046^b$	$P = 0.167^c$
EPIC urinary (minimum *n*)	745	112	124	503
Urinary summary	92.7 (9.1)	93.3 (8.1)	91.3 (10.5)	93.6 (8.2)
P value		$P = 0.619^a$	$P = 0.163^b$	$P = 0.205^c$
Urinary function	95.1 (8.4)	95.7 (6.7)	94.4 (8.7)	96.3 (7.0)
P value		$P = 0.826^a$	$P = 0.339^b$	$P = 0.006^c$
Urinary bother	91.0 (11.7)	91.3 (11.6)	89.2 (13.6)	91.6 (10.6)
P value		$P = 0.758^a$	$P = 0.122^b$	$P = 0.987^c$
Incontinence	93.0 (11.3)	93.8 (9.6)	91.3 (12.7)	94.7 (9.8)
P value		$P = 0.782^a$	$P = 0.128^b$	$P = 0.012^c$
Irritative/obstructive	93.0 (9.2)	93.5 (8.4)	91.8 (10.4)	93.3 (8.4)
P value		$P = 0.633^a$	$P = 0.186^b$	$P = 0.982^c$
EPIC bowel (minimum *n*)	748	113	126	509
Bowel summary	93.6 (8.4)	92.8 (8.4)	92.0 (9.9)	94.0 (7.2)
P value		$P = 0.141^a$	$P = 0.024^b$	$P = 0.843^c$
Bowel function	92.0 (8.8)	91.5 (8.2)	90.5 (9.2)	92.4 (7.9)
P value		$P = 0.335^a$	$P = 0.065^b$	$P = 0.436^c$
Bowel bother	95.1 (10.0)	93.9 (11.1)	93.0 (14.1)	95.6 (8.5)
P value		$P = 0.122^a$	$P = 0.110^b$	$P = 0.958^c$

Abbreviations: ProtecT, prostate testing for cancer and treatment; QoL, quality of life; RCT, randomized controlled trial; SD, standard deviation; SF-12, Short-form 12-item questionnaire; EQ5D, Euroqol 5 dimension scale; ICIQ, International Consultation on Incontinence Questionnaire; EPIC, Expanded Prostate Cancer Index Composite.

  [a] Comparison between randomized and advanced men.
  [b] Comparison between randomized and excluded men.
  [c] Comparison between randomized and preference men.
  [d] Differences of interest when using the $\geq$0.5 SDS cutoff for continuous outcomes or the $\leq$0.9 RR cutoff for categorical outcomes.

50−69 years [18]. However, the Zelen design then prevented access to data from potential participants who did not respond to the invitation to PSA testing or would have been excluded with serious comorbidities by primary-care physicians. Those who attended for PSA testing were probably representative of healthy men aged 50−69 years likely to respond to screening, rather than all men of the same age. While this would not seriously affect the generalizability of the ProtecT RCT to men fit for radical treatments, it remains a limitation in relation to the wider range of men diagnosed with PCa in routine practice. Extending the comprehensive cohort enabled follow-up of some less-fit men and those with more advanced disease.

The prospective study of PSA testing served as a recruitment framework for the ProtecT treatment RCT, and baseline data collected during testing and PCa diagnosis allowed the exploration of response and clinical factors that might affect the generalizability of the RCT findings. Very few differences were evident between eligible and ineligible groups (other than expected clinical factors), although men who declined diagnostic tests were slightly more materially deprived than consenters. Similarly, in the comprehensive cohort, men who chose treatments ("treatment-choice" group) were more likely to be in professional occupations and less materially deprived than those agreeing to randomization, but were otherwise almost identical. In the extended comprehensive cohort, socio-demographic characteristics were very similar between the randomized and "advanced cancer" or "excluded-other" groups, but the groups were different clinically, representing a wider range of patients who would receive the RCT treatments or other approaches such as hormone therapy in routine practice.

This study had several strengths and limitations. Many RCTs fail to include sufficient numbers of older people, women, ethnic minorities, and those with greater deprivation [28]. The prospective study recruited in areas outside London with very small numbers from ethnic minorities, and participation rates were proportionate to those populations [27]. The lack of diversity is a limitation in terms of wider representativeness, although treatment outcomes have recently been shown to be similar between ethnic groups in the United States [29]. More deprived individuals were less likely to respond at each stage in the prospective study, suggesting that new approaches to encourage participation of these groups are required.

Recruitment is challenging for many RCTs [30]. It has been suggested that pragmatic RCTs requiring "usual care" comparators should be embedded in prospective cohort studies in which participants have already consented to take part: "cohort multiple RCTs" [31]. This design could allow many RCTs to be conducted, although only among those agreeing to multiple study participation and thus not addressing important issues of response bias. Other design solutions to recruitment difficulties include "preference" designs where intervention preferences are elicited, and those without strong views are randomized [32], or comprehensive

cohorts, where those who decline randomization and choose a treatment are also followed-up alongside those randomized [22]. Each design raises ethical and practical issues, but successful examples [33] indicate that they can produce considerable data to assess generalizability and so should be further explored practically and methodologically.

Baseline characteristics were well balanced between the arms in both CAP [18] and ProtecT [19] RCTs, with high levels of retention and follow-up indicative of good internal validity and the robustness of the findings. Another strength of ProtecT was the high level of randomization of eligible participants—at 62%, much higher than the similar Prostatectomy versus Observation Trial (14.6%) [17] and most other cancer RCTs [34]. This was achieved by the integration of qualitative research to optimize recruitment and informed consent [21,35] and dedicated staff training [36]. Men who declined randomization were very similar to the randomized in almost every aspect, except for having more professional occupations and lower deprivation. These patients, choosing treatments as they would in usual care, along with the extended comprehensive cohort of patients excluded from the RCT but likely to be encountered in routine care, will provide many opportunities for analysis of clinical relevance and generalizability in due course.

It will be important to assess the impact of changes in PCa diagnosis since recruitment, such as the introduction of multi-parametric magnetic resonance imaging [37]. There have also been changes in treatment techniques, including robot-assisted surgery and developments in radiotherapy and methods of active surveillance, although recent evaluations of short- and medium-term oncological and patient-reported functional outcomes—expected to be better with newer techniques—have produced remarkably similar results to ProtecT [38−40], suggesting ProtecT's continuing clinical relevance. A recent English national audit showed that the majority of patients receiving surgery in 2014−2015 had a much higher grade and stage profile than those in ProtecT [41]—but the audit included men diagnosed clinically with symptoms as well from low-background PSA testing (around 6% p.a. in the UK [42]).

The small number of men with high-risk PCa randomized in ProtecT is a limitation, but some of these were included in the extended comprehensive cohort "advanced" and "excluded" groups, followed-up observationally [43]. The prospective study suggested that other high-risk men had moved earlier into routine care through previous PSA testing or urological treatment (Table 1), and others would be among those who declined the PSA-test invitation at the outset. The ProtecT findings are likely to be most relevant for men with low- and intermediate-risk PCa and fit for treatment, who represent a large proportion of cases diagnosed in the UK, and even higher proportions in many parts of Europe and North America where higher levels of PSA testing occur.

The PRECIS-2 tool provides support for trialists to discuss intentions to be more or less pragmatic and be clearer about the influence of design choices on

applicability [6], but some have suggested that, although useful, this is only the first stage, and that operational challenges (and solutions) during trial conduct can have a greater impact on a trial's generalizability [10,13]. The ProtecT and CAP RCTs were designed long before these tools/guidelines were available. Although they would undoubtedly have been helpful, we would suggest that alongside these tools/guidelines at the design stage, trialists should also ensure that they collect robust data that will later permit evidence-based insights into generalizability.

Such design decisions will inevitably have an impact on resources, but adding comprehensive-cohort studies and extending them as in this study are likely to provide considerable added value at relatively little cost. A real or perceived lack of relevance to patients in routine care continues to contribute to the slow or limited uptake of RCT evidence into practice [4,5]. Pragmatic RCTs need to provide applicable evidence, and initiatives such as those reported in this study are needed to provide evidence to increase clinicians' and policy-makers' confidence in the generalizability of trial findings.

5. Conclusions

Even the most pragmatic RCTs have limitations in terms of generalizability, but it is usually difficult to determine whether these relate to decisions made at the time of design, during trial conduct, or because of changes in clinical practice. Although some of these issues may be mitigated through the use of tools such as PRECIS-2 or guidelines when making design decisions, the scale and scope of most pragmatic RCTs mean that unanticipated limitations will arise before the outcomes are published. The assessment of the generalizability of the findings of an RCT requires wide-ranging judgments about the design and conduct of the RCT, the characteristics of its participants, and the relevance of the interventions and outcomes in the context of contemporary clinical practice. The embedding of the ProtecT RCT in a prospective study and with an extended comprehensive cohort, enabled data to be collected to support evidence-based judgments by clinicians and policy-makers about the generalizability of the randomized outcomes to patients in routine practice. With increasing willingness to undertake pragmatic RCTs to inform policy and practice, and tools to assist design and implementation, attention now also needs to be devoted to ensuring the collection of data that will provide insights into the generalizability of the randomized findings and facilitate the application of evidence more easily into clinical practice.

Acknowledgments

Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.jclinepi.2017.12.019.

References

[1] Rothwell PM. External validity of randomised controlled trials: "to whom do the results of this trial apply?". Lancet 2005; 365:82–93.

[2] Kennedy-Martin T, Curtis S, Faries D, Robinson S, Johnston J. A literature review on the representativeness of randomized controlled trial samples and implications for the external validity of trial results. Trials 2015;16:495.

[3] Dekkers OM, von Elm E, Algra A, Romijn JA, Vandenbroucke JP. How to assess the external validity of therapeutic trials: a conceptual approach. Int J Epidemiol 2010;39:89–94.

[4] Rothwell PM. Commentary: external validity of results of randomized trials: disentangling a complex concept. Int J Epidemiol 2010; 39:94–6.

[5] 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.

[6] Loudon K, Treweek S, Sullivan F, Donnan P, Thorpe KE, Zwarenstein M. The PRECIS-2 tool: designing trials that are fit for purpose. BMJ 2015;350:h2147.

[7] Dekkers OM, Bossuyt PM, Vandenbroucke JP. How trial results are intended to be used: is PRECIS-2 a step forward? J Clin Epidemiol 2017;84:25–6.

[8] Groenwold RHH, Dekkers OM. Designing pragmatic trials-what can we learn from lessons learned? J Clin Epidemiol 2017;90:3–5.

[9] Riddle DL. Consequences of randomized clinical trial design decisions need to be clarified. J Clin Epidemiol 2016;77:13–4.

[10] Zuidgeest MG, Goetz I, Grobbee DE, Consortium WPotG. PRECIS-2 in perspective: what is next for pragmatic trials? J Clin Epidemiol 2017;84:22–4.

[11] Zwarenstein M, Treweek S, Loudon K. PRECIS-2 helps researchers design more applicable RCTs while CONSORT Extension for Pragmatic Trials helps knowledge users decide whether to apply them. J Clin Epidemiol 2017;84:27–9.

[12] Oude Rengerink K, Kalkman S, Collier S, Ciaglia A, Worsley SD, Lightbourne A, et al. Participant eligibility, recruitment, and retention in pragmatic trials. J Clin Epidemiol 2017;89:173–80.

[13] Zuidgeest MGP, Goetz I, Groenwold RHH, Irving E, van Thiel G, Grobbee DE, et al. Pragmatic trials and real-world evidence-an introduction to the series. J Clin Epidemiol 2017;88:7–13.

[14] Andriole GL, Crawford ED, Grubb RL 3rd, Buys SS, Chia D, Church TR, et al. Prostate cancer screening in the randomized Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial: mortality results after 13 years of follow-up. J Natl Cancer Inst 2012;104:125–32.

[15] Holmberg L, Bill-Axelson A, Helgesen F, Salo JO, Folmerz P, Haggman M, et al. A randomized trial comparing radical prostatectomy with watchful waiting in early prostate cancer. N Engl J Med 2002;347:781–9.

[16] Schroder FH, Hugosson J, Roobol MJ, Tammela TL, Zappa M, Nelen V, et al. Screening and prostate cancer mortality: results of the European randomised study of screening for prostate cancer (ERSPC) at 13 years of follow-up. Lancet 2014;384:2027–35.

[17] Wilt TJ, Brawer MK, Jones KM, Barry MJ, Aronson WJ, Fox S, et al. Radical prostatectomy versus observation for localized prostate cancer. N Engl J Med 2012;367:203–13.

[18] Turner EL, Metcalfe C, Donovan JL, Noble S, Sterne JA, Lane JA, et al. Design and preliminary recruitment results of the Cluster randomised triAl of PSA testing for Prostate cancer (CAP). Br J Cancer 2014;110:2829–36.

[19] Lane JA, Donovan JL, Davis M, Walsh E, Dedman D, Down L, et al. Active monitoring, radical prostatectomy, or radiotherapy for localised prostate cancer: study design and diagnostic and baseline results of the ProtecT randomised phase 3 trial. Lancet Oncol 2014;15:1109–18.

[20] Donovan J, Hamdy F, Neal D, Peters T, Oliver S, Brindle L, et al. Prostate testing for cancer and treatment (ProtecT) feasibility study. Health Technol Assess 2003;7:1–88.

[21] Donovan JL, Lane JA, Peters TJ, Brindle L, Salter E, Gillatt D, et al. Development of a complex intervention improved randomization and informed consent in a randomized controlled trial. J Clin Epidemiol 2009;62:29–36.

[22] Schmoor C, Olschewski M, Schumacher M. Randomized and non-randomized patients in clinical trials: experiences with comprehensive cohort studies. Stat Med 1996;15:263–71.

[23] Hamdy FC, Donovan JL, Lane JA, Mason M, Metcalfe C, Holding P, et al. 10-Year outcomes after monitoring, surgery, or radiotherapy for localized prostate cancer. N Engl J Med 2016;375:1415–24.

[24] Donovan JL, Hamdy FC, Lane JA, Mason M, Metcalfe C, Walsh E, et al. Patient-reported outcomes after monitoring, surgery, or radiotherapy for prostate cancer. N Engl J Med 2016;375:1425–37.

[25] Down L, Metcalfe C, Avery K, Noble S, Lane JA, Neal DE, et al. Factors distinguishing general practitioners who more readily participated in a large randomized trial were identified. J Clin Epidemiol 2009;62:67–73.

[26] Walsh E, Turner E, Lane A, Donovan J, Neal D, Hamdy F, et al. Characteristics of men responding to an invitation to undergo testing for prostate cancer as part of a randomized trial. Trials 2016;17:497.

[27] Lane JA, Metcalfe C, Young GJ, Peters TJ, Blazeby J, Avery K, et al. Patient-reported outcomes in the ProtecT randomised trial of clinically localised prostate cancer treatments: design and baseline urinary, bowel and sexual function and quality of life. BJU Int 2016;118:869–79.

[28] Bartlett C, Doyal L, Ebrahim S, Davey P, Bachmann M, Egger M, et al. The causes and effects of socio-demographic exclusions from clinical trials. Health Technol Assess 2005;9. https://dx.doi.org/10.3310/hta9380.

[29] Wilt TJ, Jones KM, Barry MJ, Andriole GL, Culkin D, Wheeler T, et al. Follow-up of prostatectomy versus observation for early prostate cancer. N Engl J Med 2017;377:132–42.

[30] McDonald AM, Knight RC, Campbell MK, Entwistle VA, Grant AM, Cook JA, et al. What influences recruitment to randomised controlled trials? A review of trials funded by two UK funding agencies. Trials 2006;7:9.

[31] Relton C, Torgerson D, O'Cathain A, Nicholl J. Rethinking pragmatic randomised controlled trials: introducing the "cohort multiple randomised controlled trial" design. BMJ 2010;340:c1066.

[32] Brewin CR, Bradley C. Patient preferences and randomised clinical trials. BMJ 1989;299:313–5.

[33] Grant AM, Cotton SC, Boachie C, Ramsay CR, Krukowski ZH, Heading RC, et al. Minimal access surgery compared with medical management for gastro-oesophageal reflux disease: five year follow-up of a randomised controlled trial (REFLUX). BMJ 2013;346:f1908.

[34] Mc Daid C, Hodges Z, Fayter D, Stirk L, Eastwood A. Increasing participation of cancer patients in randomised controlled trials: a systematic review. Trials 2006;7:16.

[35] Donovan JL, Rooshenas L, Jepson M, Elliott D, Wade J, Avery K, et al. Optimising recruitment and informed consent in randomised controlled trials: the development and implementation of the Quintet Recruitment Intervention (QRI). Trials 2016;17(1):283.

[36] Lane JA, Wade J, Down L, Bonnington S, Holding PN, Lennon T, et al. A Peer Review Intervention for Monitoring and Evaluating sites (PRIME) that improved randomized controlled trial conduct and performance. J Clin Epidemiol 2011;64:628–36.

[37] Ahmed HU, El-Shater Bosaily A, Brown LC, Gabe R, Kaplan R, Parmar MK, et al. Diagnostic accuracy of multi-parametric MRI and TRUS biopsy in prostate cancer (PROMIS): a paired validating confirmatory study. Lancet 2017;389:815–22.

[38] Yaxley JW, Coughlin GD, Chambers SK, Occhipinti S, Samaratunga H, Zajdlewicz L, et al. Robot-assisted laparoscopic prostatectomy versus open radical retropubic prostatectomy: early outcomes from a randomised controlled phase 3 study. Lancet 2016;388:1057–66.

[39] Barocas DA, Alvarez J, Resnick MJ, Koyama T, Hoffman KE, Tyson MD, et al. Association between radiation therapy, surgery, or observation for localized prostate cancer and patient-reported outcomes after 3 years. JAMA 2017;317:1126–40.

[40] Chen RC, Basak R, Meyer A, Kuo T, Carpenter WR, Agans RP, et al. Association between choice of radical prostatectomy, external beam radiotherapy, brachytherapy, or active surveillance and patient-reported quality of life among men with localized prostate cancer. JAMA 2017;317:1141–50.

[41] English National Prostatectomy Audit. Available at http://www.npca.org.uk/wpcontent/uploads/2016/12/NPCA-2016-Annual-Report-Final.pdf. Accessed February 23, 2017.

[42] Williams N, Hughes LJ, Turner EL, Donovan JL, Hamdy FC, Neal DE, et al. Prostate-specific antigen testing rates remain low in UK general practice: a cross-sectional study in six English cities. BJU Int 2011;108:1402–8.

[43] Johnston TJ, Shaw GL, Lamb AD, Parashar D, Greenberg D, Xiong T, et al. Mortality among men with advanced prostate cancer excluded from the ProtecT trial. Eur Urol 2017;71:381–8.