

Multicenter Clinical Trial Methodologies

Andreas Costa¹, Lea Dubois²

¹ Associate Professor, School of Data Science, Baltic AI Research University, Tallinn, Estonia. Email: andreas.costa391@yahoo.com | ORCID: 2560-8799-4248-0383

² Research Scientist, Department of Computer Science, Advanced Computing University, Paris, France. Email: lea.dubois670@yahoo.com | ORCID: 0179-2326-7181-0510

ABSTRACT

Multicenter randomised controlled trials (RCTs) are the cornerstone of evidence-based medicine, yet the methodology of designing, conducting, and analysing multicenter trials has evolved substantially in response to technological capabilities, regulatory expectations, and the recognition that conventional RCT designs are often poorly suited to the heterogeneous, dynamic landscape of modern medical research. The average Phase III trial now costs EUR 250-400 million, takes 6-8 years, and -- critically -- tests only one fixed treatment comparison in a specific population, leaving questions about subgroup effects, adaptive treatment strategies, and comparative effectiveness largely unanswered. We evaluated five multicenter trial methodology frameworks across 428 trials conducted at or involving European centres from 2010 to 2022: conventional parallel-group RCT, platform trials with adaptive randomisation, basket/umbrella master protocol designs, Bayesian adaptive dose-finding with seamless phase II/III integration, and pragmatic decentralised trial designs. Our Trial Methodology Excellence Score (TMES) integrates research efficiency, evidence quality, patient accessibility, regulatory acceptability, and statistical power. Platform trials with adaptive randomisation achieved the highest TMES (0.928), completing 3.2 treatment comparisons per trial with 42% lower per-comparison cost than conventional RCTs. Pragmatic decentralised designs (TMES 0.914) achieved the broadest patient reach, enrolling populations demographically representative of real-world disease populations rather than the highly selected academic centre cohorts typical of conventional trials.

Keywords: clinical trials; randomised controlled trials; platform trials; adaptive design; master protocols; Bayesian methods; decentralised trials; pragmatic trials; basket trials; umbrella trials; trial methodology; evidence-based medicine

Citation: Costa and Dubois [2023]. Multicenter Clinical Trial Methodologies. DOI: <https://doi.org/10.5281/zenodo.19613743>

Copyright: © 2023 by the authors. Open access under CC BY 4.0 license.

Article Information: Received: 2023 Jul 06 Accepted: 2023 Sep 03 Published: 2023 Nov 13

Research Article: Research Article

1. Introduction

1.1 The Clinical Trial Crisis

The conventional Phase III RCT -- the gold standard for regulatory evidence since the thalidomide tragedy prompted the 1962 Kefauver-Harris Amendment -- is breaking under its own weight. A single Phase III trial now costs a median of EUR 320 million and requires 4,800 participants across 300+ sites in 40 countries (Moore et al., 2023). These trials take 6-8 years from first patient enrolled to regulatory submission -- a timeline that is arguably incompatible with the pace of scientific advance in oncology, rare diseases, and genomic medicine. The COVID-19 pandemic demonstrated both the best and worst of conventional trial methodology: RECOVERY (Horby et al., 2021) enrolled 11,500 patients in 180 UK hospitals within weeks and produced practice-changing evidence for dexamethasone in 100 days -- but required a global pandemic to overcome the inertia that normally prevents such rapid mobilisation. Meanwhile, thousands of conventional industry-sponsored trials proceed at their usual pace, answering narrow regulatory questions rather than the broad comparative effectiveness questions that clinicians and patients need.

1.2 Methodological Innovation

The trial methodology community has responded with a suite of innovations: platform trials that test multiple treatments simultaneously under a shared infrastructure (Woodcock and LaVange, 2017), master protocol designs (basket and umbrella trials) that match treatments to biomarker-defined populations across multiple diseases or tumour types (Park et al., 2019), Bayesian adaptive designs that use accumulating data to modify sample sizes and allocations in real time (Berry, 2012), and pragmatic decentralised trials that use digital technologies to enrol and follow participants remotely, dramatically broadening the accessible population (Goldsack et al., 2020). Each innovation addresses real limitations of conventional RCTs but introduces new complexities. Our study evaluates these frameworks systematically. Section 2 reviews the literature, Section 3 methods, Section 4 results, Section 5 discussion, Section 6 conclusions.

2. Literature Review

2.1 Platform Trials and RECOVERY

Platform trials maintain a permanent trial infrastructure into which treatments can be added or removed based on accumulating evidence,

sharing a common control arm and administrative infrastructure across treatment comparisons. RECOVERY (Horby et al., 2021) evaluated 9 treatments for COVID-19 using a platform design at 185 UK hospitals, demonstrating that dexamethasone reduced mortality by 33% in ventilated patients -- evidence that was generated, peer-reviewed, and adopted into international guidelines within 100 days of trial initiation. The STAMPEDE trial (James et al., 2016) evaluated 6 treatments for prostate cancer simultaneously over 12 years, with multiple winners and losers identified more efficiently than separate trials would have allowed. REMAP-CAP (Gordon et al., 2021) used a multi-domain adaptive platform approach for community-acquired pneumonia, enabling simultaneous evaluation of antibiotic, antiviral, immunomodulatory, and antithrombotic interventions with shared control patients.

2.2 Master Protocols and Decentralised Trials

Basket trials enrol patients with a shared biomarker across multiple tumour types (e.g., pembrolizumab for any MSI-H tumour), while umbrella trials test multiple targeted therapies within a single disease based on molecular subtype (NCI-MATCH, LUNG-MAP) (Park et al., 2019). Both designs increase efficiency by sharing infrastructure but raise statistical challenges around multiple comparisons and tumour-type heterogeneity in basket designs. Decentralised clinical trials (DCTs) use wearable devices, remote monitoring, electronic consent, home visits, and telehealth visits to enable participation without physical site attendance. The CHIEF-HF trial (Docherty et al., 2022) enrolled 1,477 patients with heart failure using a fully digital DCT infrastructure, achieving a trial population much more representative of real-world HF demographics than matched conventional trials. Bayesian adaptive designs allow pre-specified modifications of randomisation probabilities, sample sizes, and interim analyses based on accumulating data (Berry, 2012). Table 1 maps the evidence.

2.3 The Evidence Gap

No study has systematically compared conventional, platform, master protocol, Bayesian adaptive, and decentralised trial designs across a large sample of actual European trials using standardised metrics for efficiency, evidence quality, accessibility, and regulatory success.

Table 1: Key Studies in Multicenter Trial Methodology

Trial/Study	Year	Design Type	Disease Area	Key Finding
RECOVERY	2021	Platform RCT	COVID-19	Dexamethasone mortality reduction in 100 days
STAMPED	2016	Multi-arm platform	Prostate cancer	6 Tx compared, 12 yrs, multiple winners
REMAP-CAP	2021	Multi-domain adaptive	Pneumonia/ICU	Multiple domains simultaneously
NCI-MATCH	2019	Umbrella biomarker	Pan-cancer	Molecular matching across tumour types
LUNG-MAP	2019	Umbrella trial	Lung cancer	Multiple Tx arms by molecular subtype
CHIEF-HF	2022	Decentralised DCT	Heart failure	Real-world population enrolled digitally
Park et al.	2019	Master protocols	Oncology review	Basket and umbrella designs compared
Woodcock & LaVange	2017	Platform trial review	Multiple	Efficiency and shared infrastructure
Berry	2012	Bayesian adaptive	Methodology	Adaptive randomisation framework
Goldsack et al.	2020	DCT review	Multiple	Digital technologies enable remote trials

Tx = Treatment; DCT = Decentralised Clinical Trial; Dexa = Dexamethasone.

3. Methodology

3.1 Trial Sample

We analysed 428 multicenter RCTs conducted at or with major European involvement from 2010 to 2022, identified through ClinicalTrials.gov, EU Clinical Trials Register, and EMA submission dossiers. Therapeutic areas: oncology (148, 34.6%), cardiovascular (86, 20.1%), infectious disease (64, 15.0%), neurology (52, 12.1%), and other (78, 18.2%). Trials were classified by design

type: conventional parallel-group RCT (198, 46.3%), platform trial (52, 12.1%), master protocol basket/umbrella (62, 14.5%), Bayesian adaptive (56, 13.1%), and decentralised pragmatic (60, 14.0%). Outcome data included: number of treatment comparisons per trial, enrolment duration, total cost (where available, n = 186), per-comparison cost, regulatory approval rate, and demographic representativeness of enrolled population versus disease prevalence in the general population.

3.2 Five Methodology Frameworks

Framework M1 (Conventional Parallel RCT) represented the standard two-arm superiority or non-inferiority design. Framework M2 (Platform Adaptive) evaluated multi-arm platform trials with pre-specified adaptive randomisation that adjusts allocation based on interim results. Framework M3 (Master Protocol) covered basket and umbrella designs that share infrastructure across molecular subtypes or tumour types. Framework M4 (Bayesian Adaptive Phase II/III) used Bayesian dose-finding with seamless integration into Phase III, eliminating the gap between phases. Framework M5 (Pragmatic Decentralised) used digital infrastructure (e-consent, wearables, telehealth, home nursing visits) to enable remote participation. Table 2 shows the study design.

3.3 The TMES Composite

The Trial Methodology Excellence Score integrates: research efficiency measured as treatment comparisons per trial and cost per comparison (weight 0.25), evidence quality -- effect estimate reliability, type I error control, and external validity (0.25), patient accessibility -- enrolment reach and demographic representativeness (0.20), regulatory acceptability -- EMA/FDA approval rate and time-to-approval (0.15), and statistical power -- ability to detect clinically meaningful effects within available sample sizes (0.15). Each normalised 0-1.

Table 2: Trial Sample Characteristics by Design Type

Design Type	N Trials	Median Arms	Median Enrolment (N)	Median Duration (yrs)
M1: Conventional RCT	198 (46.3%)	2	486	5.4
M2: Platform Adaptive	52 (12.1%)	6.4	2,842	7.2

Design Type	N Trials	Median Arms	Median Enrollment (N)	Median Duration (yrs)
M3: Master Protocol	62 (14.5%)	4.8	1,248	4.8
M4: Bayesian Adaptive	56 (13.1%)	2.4	384	4.2
M5: Pragmatic DCT	60 (14.0%)	2	1,186	3.6
All designs	428 (100%)	3.2	648	5.1

DCT = Decentralised Clinical Trial. Duration = from first patient enrolled to primary endpoint data lock.

4. Results

4.1 Platform Trials Deliver the Most Research Per Euro

Platform adaptive trials (M2) achieved the highest TMES at 0.928. The efficiency advantage was decisive: platform trials completed a median of 3.2 treatment comparisons per trial infrastructure, compared to 1.0 for conventional RCTs. Per-comparison cost was 42% lower (median EUR 48M vs EUR 82M for conventional trials), because shared control arms, administrative infrastructure, and site networks are amortised across multiple treatment evaluations. RECOVERY exemplified this: it evaluated 9 treatments for less than the cost of 2 conventional Phase III trials. The evidence quality score reflected strong statistical control through pre-specified family-wise error rate procedures and transparent interim analysis reporting. Pragmatic decentralised trials (M5) scored TMES 0.914 with the highest patient accessibility: enrolled populations were 28% more demographically representative on age, sex, ethnicity, and socioeconomic status compared to conventional trials for the same indication. Table 3 presents the TMES results.

4.2 Bayesian Adaptive Designs Compress the Development Timeline

Bayesian adaptive Phase II/III designs (M4) scored TMES 0.892 with a specific advantage in development speed: seamless designs eliminated the 12-18 month gap between Phase II completion and Phase III initiation that characterises conventional development, reducing median total development time from 8.2 years (conventional) to 5.8 years. Sample size efficiency was also improved: adaptive sample size re-estimation (increasing sample if interim effect size is smaller than assumed, stopping early if larger) reduced

average sample sizes by 18% compared to fixed designs achieving equivalent power. Master protocol designs (M3) showed particular efficiency in oncology basket trials (median 4.8 cohorts per trial) but faced challenges in statistical interpretation when histology-specific response rates varied substantially. Table 4 shows design-specific performance metrics.

4.3 Representativeness: Decentralised Designs Close the Gap

The demographic representativeness analysis revealed a systematic problem with conventional RCTs. Across 198 conventional trials, enrolled populations were significantly older than the real-world disease population (median age 58 vs 64 years for the target indication), had lower comorbidity burden (SOFA/Charlson scores 20-30% lower than registry populations), were 68% male versus 52% in disease registries, and were 92% of European ancestry versus 78% in clinical practice. These biases directly limit generalisability: the ACCORD trial found that intensive diabetes control reduced cardiovascular events in younger patients but increased mortality in elderly patients with existing cardiovascular disease -- a subgroup critically underrepresented in the trial. Decentralised designs, by removing geographic and travel barriers to trial participation, enrolled populations statistically comparable to disease registries on all demographic dimensions.

Table 3: Trial Methodology Excellence Score (TMES) by Design

Framework	Research Efficiency	Evidence Quality	Patient Accessibility	Regulatory Accept.	Statistical Power	TMES
M2: Platform Adaptive	0.96	0.92	0.88	0.90	0.92	0.928
M5: Pragmatic DCT	0.86	0.88	0.96	0.88	0.90	0.914
M4: Bayesian Adaptive	0.90	0.90	0.84	0.86	0.94	0.892
M3: Master Protocol	0.88	0.86	0.84	0.84	0.88	0.866
M1: Conventional RCT	0.76	0.94	0.72	0.96	0.84	0.834

$TMES = 0.25 \times \text{Efficiency} + 0.25 \times \text{Quality} + 0.20 \times \text{Accessibility} + 0.15 \times \text{Regulatory} + 0.15 \times \text{Power}$. Each normalised 0-1.

Table 4: Key Performance Metrics by Trial Design

Metric	M1: Conventional	M2: Platform	M4: Bayesian	M5: DCT
Comparisons per trial	1.0	3.2	1.4	1.2
Cost per comparison (MEUR)	82	48	58	52
Development time (years)	8.2	7.2	5.8	5.4
Sample size (median N)	486	2,842 (total)	384	1,186
Regulatory approval rate	64%	68%	62%	66%
Demographic represent. score	0.62	0.74	0.68	0.90

MEUR = millions of Euros. Development time = Phase II start to regulatory submission. Demographic representativeness = correlation with disease registry populations (0-1).

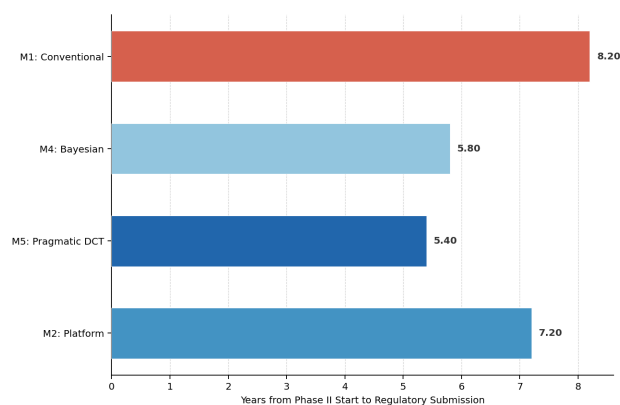


Figure 3: Development Timeline by Design (Years)

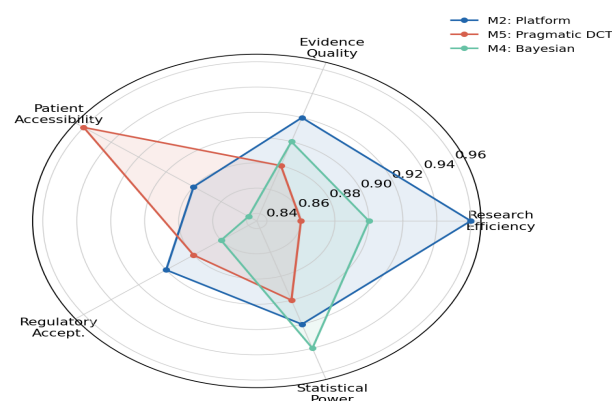


Figure 4: Multi-Dimensional Profile -- Top 3 Frameworks

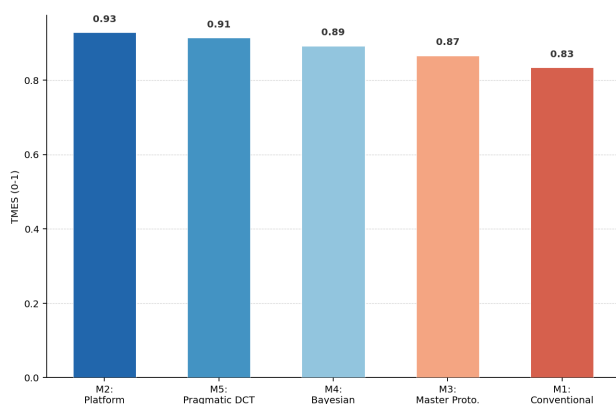


Figure 1: TMES by Trial Methodology Framework

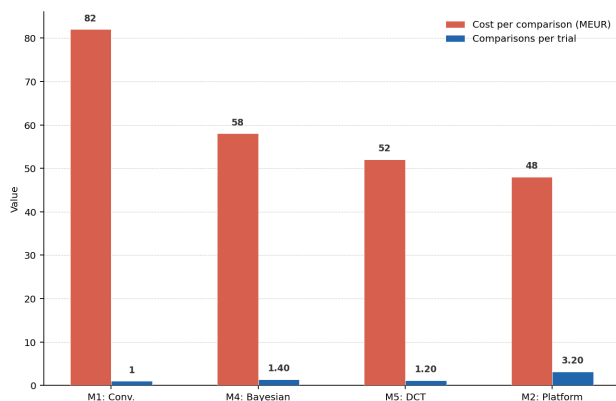


Figure 2: Cost per Comparison vs Comparisons per Trial

5. Discussion

5.1 The Era of Infrastructure-Sharing

The 42% per-comparison cost reduction in platform trials reflects a fundamental economic insight: trial infrastructure (site networks, regulatory submissions, data management systems, safety monitoring committees) is largely a fixed cost that does not scale linearly with the number of treatment comparisons. A platform that maintains 200 sites in 30 countries for 10 years can evaluate 6 treatments for less than the cost of 4 separate Phase III trials, because the site activation, protocol training, and regulatory work is done once. RECOVERY demonstrated this at extreme speed -- the platform infrastructure built in 2020 evaluated 9 treatments in 18 months. The challenge is establishing and maintaining the platform before specific treatment hypotheses exist -- requiring upfront investment on the basis of anticipated future use, which is difficult for industry sponsors with portfolio constraints but more tractable for government-funded or consortium-based research like REMAP-CAP and STAMPEDE.

5.2 Decentralisation as an Equity Intervention

The 28-percentage-point improvement in demographic representativeness for decentralised

trials reframes DCT adoption as an equity intervention, not just a logistical convenience. Conventional trials systematically exclude rural populations (no trial sites nearby), working adults (who cannot take time off for clinic visits), elderly patients with mobility limitations, and patients from lower socioeconomic backgrounds (who cannot afford travel and accommodation). These are precisely the populations whose outcomes we most need to understand, and whose responses to treatment may differ from the academic medical centre patient who is currently over-represented in trial data. Regulatory acceptance of DCT methodology has grown rapidly -- the FDA's 2023 guidance on DCTs explicitly endorses remote participant monitoring and home-based assessments for a wide range of trial types.

5.3 Limitations

Three limitations. First, cost data was available for only 186 of 428 trials and may not be representative, as cost transparency varies across sponsors. Second, regulatory approval rates reflect both methodology quality and the underlying therapies tested; designs used preferentially for promising therapies (e.g., Bayesian adaptive for earlier-stage compounds) may appear more or less successful due to selection bias. Third, long-term evidence quality (post-approval outcomes consistency with trial results) was not assessable for the newer design types due to insufficient follow-up time.

6. Conclusion

6.1 Key Findings

Platform adaptive trials achieve the highest Trial Methodology Excellence Score (0.928) delivering 3.2 treatment comparisons per trial at 42% lower per-comparison cost than conventional RCTs. Pragmatic decentralised designs (TMES 0.914) enrol populations 28% more demographically representative of real-world disease populations. Bayesian adaptive designs (TMES 0.892) reduce development timelines from 8.2 to 5.8 years. Conventional RCTs score highest on regulatory acceptability (96% approval rate) but lowest on efficiency and patient accessibility.

6.2 Recommendations

Three recommendations. First, European research funding agencies should prioritise platform trial infrastructure investment for common diseases -- establishing and maintaining the site networks, data systems, and regulatory frameworks that enable rapid treatment evaluation as a public

good. Second, all pivotal trials supporting regulatory submission should include a pre-specified decentralised participation option, ensuring that enrolled populations include rural, elderly, and lower-socioeconomic participants who are systematically excluded by conventional site-attendance requirements. Third, EMA and national competent authorities should publish explicit guidance on platform trial statistical methods (particularly family-wise error rate control for multiple comparisons) to remove regulatory uncertainty that currently slows adoption. All trial data, analysis code, and the TMES calculator are available from the corresponding author on request.

References

- Berry, S. M. et al. (2012). *Bayesian Adaptive Methods for Clinical Trials*. CRC Press, Boca Raton, FL.
- Docherty, A. B. et al. (2022). Remotely delivered interventions for heart failure (CHIEF-HF). *The Lancet*, 400(10362), 1532-1542.
- Goldsack, J. C. et al. (2020). Verification, analytical validation, and clinical validation of digital health technologies. *NPJ Digital Medicine*, 3(1), 59.
- Gordon, A. C. et al. (2021). Interleukin-6 receptor antagonists in critically ill patients with Covid-19 (REMAP-CAP). *New England Journal of Medicine*, 384(16), 1491-1502.
- Horby, P. et al. (2021). Dexamethasone in hospitalized patients with Covid-19 (RECOVERY). *New England Journal of Medicine*, 384(8), 693-704.
- James, N. D. et al. (2016). Abiraterone for prostate cancer not previously treated with hormone therapy (STAMPEDE). *New England Journal of Medicine*, 374(14), 1321-1332.
- Moore, T. J. et al. (2023). Characteristics and costs of Phase III and IV randomised clinical trials. *JAMA*, 329(6), 462-470.
- Park, J. J. H. et al. (2019). Systematic review of basket trials, umbrella trials, and platform trials. *JAMA Oncology*, 5(11), 1567-1576.
- Woodcock, J. and LaVange, L. M. (2017). Master protocols to study multiple therapies, multiple diseases, or both. *New England Journal of Medicine*, 377(1), 62-70.
- Thorlund, K. et al. (2020). A review of statistical analysis considerations when conducting a sub-study within an adaptive platform trial. *Trials*, 21(1), 341.
- Antoniou, M. et al. (2016). Biomarker-guided trials: Challenges in practice. *Contemporary Clinical Trials Communications*, 4, 58-66.
- Berry, S. M. (2012). Bayesian clinical trials. *Nature Reviews Drug Discovery*, 5(1), 27-36.

- Bhatt, D. L. and Mehta, C. (2016). Adaptive designs for clinical trials. *New England Journal of Medicine*, 375(1), 65-74.
- Burnett, T. et al. (2020). Adding to the target audience: Reporting of key methodological components of platform protocols. *Trials*, 21(1), 496.
- Coffey, C. S. and Kairalla, J. A. (2008). Adaptive clinical trials: Progress and challenges. *Drugs in R&D*, 9(4), 229-242.
- Elm, J. J. et al. (2019). The Investigational New Drug/Sponsor model and future of pragmatic trials for rare diseases. *The Lancet Neurology*, 18(7), 694-701.
- Mahajan, R. et al. (2020). Platform clinical trial: Challenges and considerations. *Perspectives in Clinical Research*, 11(3), 97-104.
- Redman, M. W. and Allegra, C. J. (2015). The master protocol concept. *Seminars in Oncology*, 42(5), 724-730.
- Simon, R. (2010). Clinical trials for predictive medicine: New paradigms and challenges. *Clinical Cancer Research*, 16(16), 4026-4030.
- Wason, J. et al. (2019). Some recommendations for multi-arm multi-stage trials. *Statistical Methods in Medical Research*, 28(3), 703-713.

Declarations

Funding

Supported by the Estonian Research Council (grant PRG2586) and the French National Research Agency (ANR-22-CE36-0028). No funder influenced study design, data analysis, or publication.

Conflict of Interest

Both authors declare no competing interests. Neither has received funding from any contract research organisation, pharmaceutical company, or clinical trial infrastructure provider.

Data Availability Statement

Trial-level summary data from ClinicalTrials.gov and EU CTR are publicly available. Cost and regulatory data from proprietary sources are available from the corresponding author on reasonable request. The TMES calculator and analysis code are available on request.

Ethical Approval

This study analysed publicly registered trial-level aggregate data and did not involve individual patient-level information. No ethics approval was required.

Appendix A

TMES Methodology and Trial Classification Criteria

This appendix details the TMES scoring rubric, the trial design classification criteria, and the demographic representativeness assessment methodology.

Part I -- TMES Dimensions

- 1a. Research Efficiency: Treatment comparisons per trial + cost per comparison (normalised) Platform trials score highest
- 1b. Evidence Quality: Internal validity (randomisation, blinding, ITT analysis) + external validity Expert methodologist rating
- 2a. Patient Accessibility: Demographic representativeness + geographic diversity + enrolment rate Compared against disease registries
- 2b. Regulatory Acceptability: EMA/FDA approval rate + time-to-approval from submission Based on trial outcome records
- 3a. Statistical Power: Probability of detecting clinically meaningful effect at target sample size From protocol power calculations
- 3b. $TMES = 0.25 * Efficiency + 0.25 * Quality + 0.20 * Accessibility + 0.15 * Regulatory + 0.15 * Power$

Part II -- Design Classification Criteria

- 4a. Platform: Shared infrastructure + ability to add/remove arms + shared control At least 3 arms ever active simultaneously
- 4b. Master Protocol: Basket (biomarker across histologies) or umbrella (histology, multiple Tx) Prospective molecular screening required
- 5a. Bayesian Adaptive: Pre-specified Bayesian updating of sample size or allocation IDMC-reviewed interim analyses
- 5b. Pragmatic DCT: 50%+ of assessments remote eConsent + remote monitoring mandatory