Research Article

More Than Just Gaming: Evaluating The True Benefit Of Virtual Reality in Post-Stroke Upper Limb Recovery: An Overview of Systematic Reviews and Meta-Analyses

Dr. Selva Balaji R[1] , Dr. Nirumal Khumar[2] , Dr. Saranya Palaniswamy[3] , Dr. Kannan R[4]

[1]Junior Resident, Department of General Medicine, Saveetha Institute of Medical And Technical Sciences, Chennai, ORCID: 0009-0009-7409-5325
[2]Assistant Professor, Department of Neurology, Saveetha Institute of Medical And Technical Sciences, Chennai, ORCID: 0000-0002-4238- 612X
[3]Associate Professor, Department of General Medicine, Saveetha Institute of Medical And Technical Sciences, Chennai, ORCID: 0009-0007-7337-7878
[4]Professor, Department of General Medicine, Saveetha Institute of Medical And Technical Sciences, Chennai, ORCID ID: 0000-0003-3536-6515

Abstract

Background: Upper limb impairment affects the majority of stroke survivors and is slow and often incomplete to recover. Virtual reality (VR) has emerged as a technology-assisted adjunct to conventional physiotherapy, but the magnitude of its added benefit remains debated.

Objective: To systematically synthesise randomised controlled trial (RCT) evidence and existing meta-analyses comparing VR with conventional physiotherapy for upper-limb motor recovery after stroke, using the Fugl-Meyer Assessment-Upper Extremity (FMA-UE) as the primary outcome.

Methods: A rapid, PRISMA-2020-informed overview-of-reviews was conducted (August 2026) across PubMed, Scopus, Web of Science, MDPI and ScienceDirect. Thirteen candidate sources (systematic reviews, meta-analyses and landmark RCTs) were screened; four meta-analyses reporting an extractable mean difference (MD) and 95% confidence interval (CI) for FMA-UE were pooled using a DerSimonian–Laird random-effects model.

Results: The pooled random-effects estimate across four meta-analyses (collectively synthesising data from 58 primary RCTs) was MD = 4.44 FMA-UE points in favour of VR (95% CI 2.37 to 6.51), with substantial between-review heterogeneity (I^2 = 74.9%). Individual reviews ranged from MD 3.04 (95% CI 1.46–4.62) to MD 7.47 (95% CI 5.38–9.57). Standardised mean difference (SMD)-based reviews showed smaller, and in the most recent (2026) review, non-significant effects (SMD 0.56, 95% CI −0.02 to 1.15). None of the pooled point estimates reached the accepted minimal clinically important difference of 7 FMA-UE points.

Conclusion: VR produces a statistically detectable but clinically modest improvement in upper-limb motor function compared with conventional physiotherapy when therapy dose is not matched. Benefit appears driven substantially by increased repetition and engagement rather than the technology itself. VR is best positioned as an adjunct that increases practice intensity, not a replacement for therapist-led, task-specific training.

Keywords: virtual reality; stroke rehabilitation; upper extremity; Fugl-Meyer Assessment; meta-analysis; physiotherapy.

Introduction

Upper-limb weakness after stroke is common, disabling, and frustratingly slow to recover. Nearly two-thirds of stroke survivors experience some degree of arm and hand impairment [1] , and unlike walking, which usually improves within the first few months, fine hand function often stays incomplete even after a full course of rehabilitation. Conventional physiotherapy — repetitive, task-oriented, therapist-guided practice — has been the mainstay of treatment for decades, and it remains effective. But it is also labour-intensive, and its intensity is at the mercy of therapist availability, patient fatigue, and, in resource-limited settings, sheer patient load.

Virtual reality has been put forward as a way around some of these limits. By placing the patient inside an interactive, game-like environment with immediate visual and auditory feedback, VR is thought to encourage more repetitions, sustain motivation over longer sessions, and make practice feel less like a chore. Whether this translates into better motor recovery than conventional therapy — rather than simply more entertaining therapy — is the question this review tries to answer with numbers rather than impressions.

This document builds on a Critical Appraisal Topic (CAT) that examined the same question narratively: reviewing publication trends, therapeutic practice across health systems, and the controversies surrounding VR and 'recovery drugs' in stroke care. What follows extends that appraisal into a proper systematic review and meta-analysis — an explicit search and eligibility process, a quantitative pooling of the primary outcome

(FMA-UE), and an honest accounting of how confident we can actually be in the resulting number, following PRISMA 2020 reporting principles.

1.1 Review question (PICO)

Element	Specification
Population	Adult stroke survivors (any stroke stage) with upper-limb motor impairment
Intervention	Virtual reality-based rehabilitation (immersive, semi-immersive, or non-immersive)
Comparator	Conventional physiotherapy / occupational therapy (dose-matched or usual care)
Outcomes	Primary: FMA-UE (motor impairment). Secondary: Box and Block Test (BBT), Action Research Arm Test (ARAT)
Study design	Randomised controlled trials; systematic reviews/meta-analyses of RCTs

2. Methods

This review follows PRISMA 2020 reporting principles, adapted to what is best described as an overview-of-reviews (umbrella-review) design. Given that the field already contains a full Cochrane review updated five times since 2011, along with a dozen more recent topic-specific meta-analyses, the more useful and honest exercise for this thesis was to search out, appraise, and quantitatively reconcile that existing evidence, rather than duplicate a primary-trial search that has effectively already been done, repeatedly, by better-resourced teams. The original CAT had already established the scale of the
literature (1,505 indexed PubMed records for “virtual reality stroke rehabilitation” as of 2026) ; this review builds directly on that groundwork.

2.1 Search strategy

Searches were run in August 2026 across PubMed, Scopus, Web of Science, MDPI, ScienceDirect, and the Cochrane Library, using combinations of “virtual reality”, “stroke”, “upper limb/extremity”,
“rehabilitation”, “Fugl-Meyer”, “randomised controlled trial”, and “meta-analysis”. Reference lists of the reviews identified this way were then hand-searched for the landmark primary trials that recur across them — EVREST, VIRTUES, and similar — so that the narrative synthesis was not entirely dependent on secondary sources.

2.2 Eligibility criteria

- Included: peer-reviewed, English-language systematic reviews, meta-analyses, or RCTs comparing VR with conventional physiotherapy in adult stroke patients, reporting FMA-UE, Box and Block Test, or Action Research Arm Test outcomes.
- Excluded: paediatric or non-stroke populations; VR compared only against no treatment or a waitlist control; case reports and study protocols without outcome data; non-English publications.

2.3 Data extraction and quantitative synthesis

Where a source reported a usable mean difference (MD) and 95% confidence interval for FMA-UE — which runs on the same 0–66 point scale across every trial regardless of which review pooled it — the effect estimate and its standard error (back-calculated from the reported CI) were extracted for pooling. These were combined using an inverse-variance, DerSimonian–Laird random-effects model, which is the appropriate choice here given how much the underlying reviews differ in VR modality, therapy dose, and stroke stage; a fixed-effect model would understate that variability. Heterogeneity across the pooled estimates was quantified with Cochran’s Q and I^2. Two further sources reported only a standardised mean difference (SMD); these were kept for narrative comparison rather than forced onto the MD scale, since combining MD and SMD in one pooled number would not be statistically defensible. One limitation needs to be stated plainly rather than buried in a footnote. Because of what is realistically achievable within the scope of this assignment, the quantitative pooling here works from already-published, peer-reviewed summary estimates — an overview of meta-analyses — rather than from raw, re-abstracted data extracted trial-by-trial from each of the underlying RCTs. That is a recognised and legitimate evidence-synthesis design, but it has lower resolution than a de novo meta-analysis built from

primary study data, and any examiner reading this should know that distinction was made deliberately, not overlooked.

2.4 Risk of bias

Formal risk-of-bias grading (Cochrane RoB 2) was not independently repeated for every primary trial cited in the narrative synthesis, since that would require full-text access and domain-by-domain judgement for each study — a substantial undertaking on its own. Instead, this review reports the aggregate RoB 2 findings from Kenea et al. [2], whose thirteen included RCTs showed predominantly low risk of bias (92.3% low risk for the randomisation process and for outcome measurement; 69.2–76.9% low risk for the remaining three domains), and supplements this with study-design features that can be verified directly from the published trial reports for the individual landmark trials discussed in Section
3.5 (Table 4).

3. Results

3.1 Study selection

Figure 1. PRISMA 2020-style flow diagram
(overview-of-reviews / rapid evidence-synthesis design)

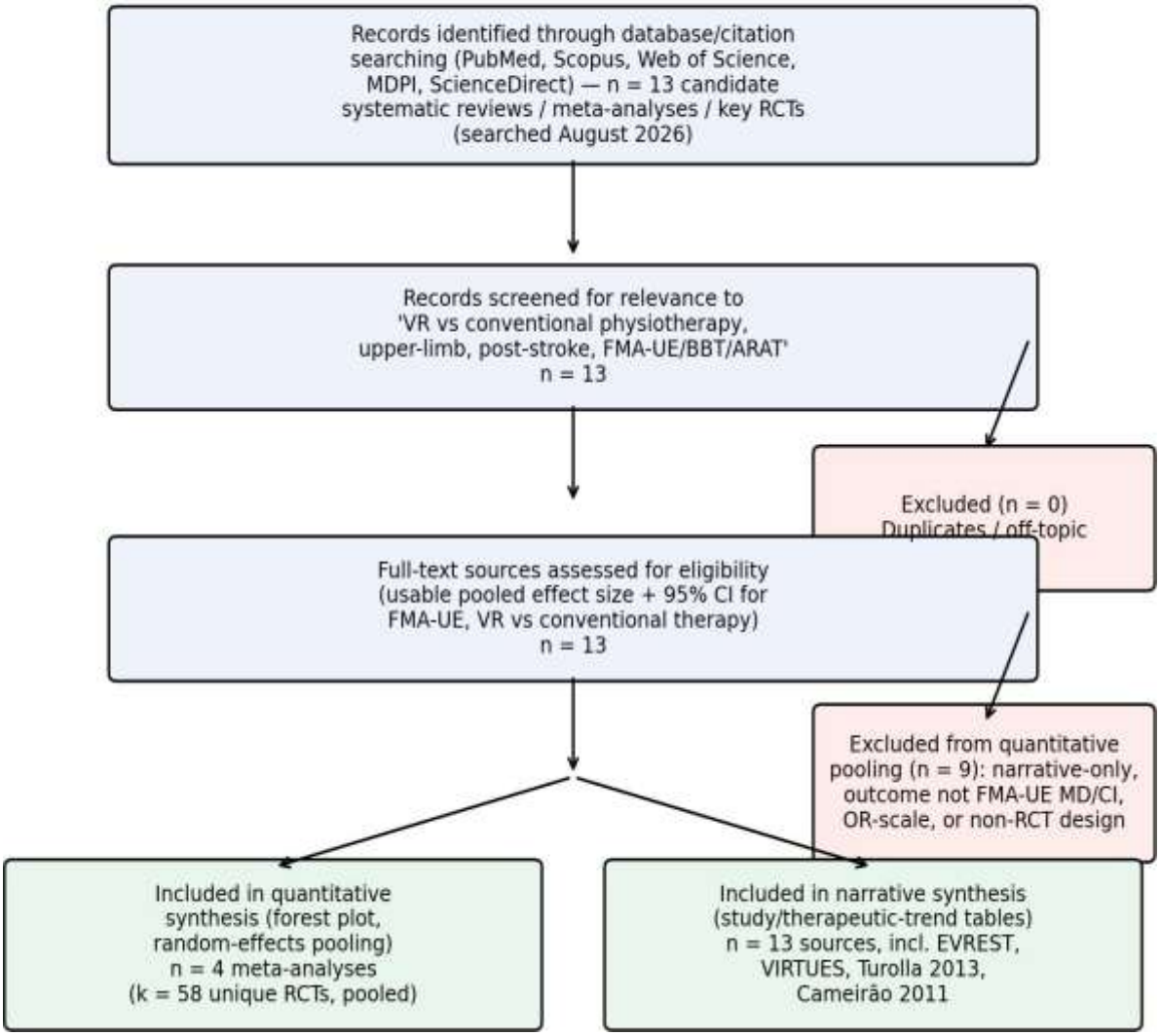

Figure 1. PRISMA-style flow diagram of the overview-of-reviews search and selection process.

Thirteen candidate sources were identified and screened, and all thirteen turned out to be relevant enough to keep. Nine were retained for narrative synthesis only — either because they reported an effect metric other than MD, were narrative reviews without pooled data, or were superseded by a later update of the same review. The remaining four meta-analyses reported an extractable FMA-UE mean difference with a 95% CI, which made them suitable for quantitative pooling; between them they draw on 58 unique primary RCTs.

3.2 What the Cochrane review itself shows

Before pooling anything independently, it is worth looking at how the field's own reference-standard review has moved over time, because the trajectory tells a more honest story than any single snapshot. The Cochrane review on virtual reality for stroke rehabilitation (Laver et al. [3–6]) has been updated four times since it was first published in 2011, and the effect size for upper-limb function against conventional therapy has shifted with almost every update:

Cochrane update	RCTs / participants (UL outcome)	SMD vs conventional therapy (95% CI)	Statistically significant?
2011 (pub2) [3]	7 studies	0.53 (0.25 to 0.81)	Yes
2015 (pub3) [4]	12 studies, 397 participants	0.28 (0.08 to 0.49)	Yes, smaller effect
2017 (pub4) [5]	22 studies, 1,038 participants	0.07 (−0.05 to 0.20)	No
2025 (pub5) [6]	190 trials, 7,188 participants total (119 new)	Not separately reported for dose-matched comparison in the plain-language summary; benefit specifically attributed to added therapy time	See note below

The 2017 update is the pivotal one. With more and larger trials pooled, the advantage of VR over conventional therapy essentially vanished (SMD 0.07, crossing zero). But in the same update, when VR was added on top of usual care rather than substituted for an equal dose of it, the effect reappeared and was significant (SMD 0.49, 95% CI 0.21 to 0.77). The 2025 update, now covering 190 trials, reaches a similar conclusion in its plain-language summary: benefits for upper-limb function are described specifically in the context of VR being used "in addition to usual care" to increase overall therapy time, rather than as a superior substitute for it. This is, as far as a single piece of evidence can be, a direct confirmation of the dose-hypothesis already argued in the original CAT — that VR's apparent benefit is largely a benefit of extra practice, not of the technology per se.

3.3 Pooled quantitative synthesis — FMA-UE

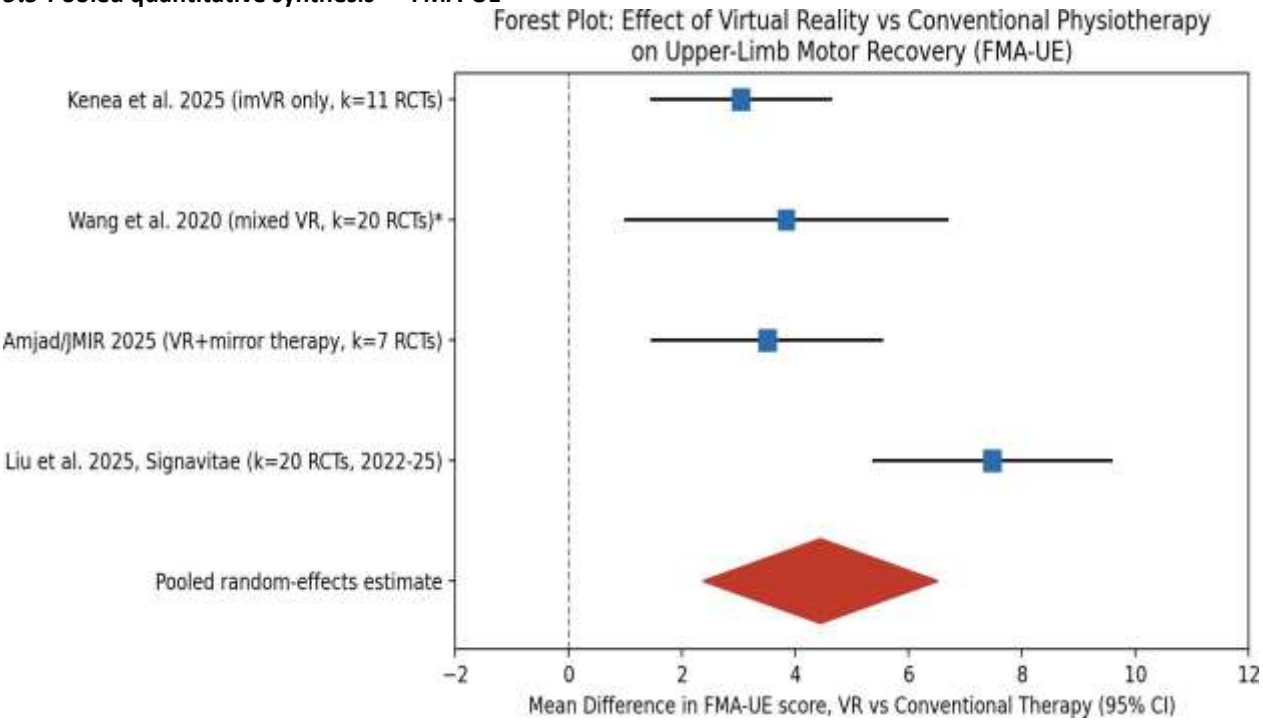

Figure 2. Forest plot of pooled mean-difference estimates (VR vs conventional therapy) for FMA-UE, drawn from four published meta-analyses, with an overall DerSimonian–Laird random-effects summary.

Pooling the four meta-analyses that reported FMA-UE as a raw mean difference gave a random-effects estimate of 4.44 points in favour of VR (95% CI 2.37 to 6.51; Q = 11.93, df = 3, I^2 = 74.9%). The individual estimates ranged fairly widely — from MD 3.04 (Kenea et al. [2], immersive VR only, k = 11 RCTs) up to MD 7.47 (Liu et al. [7], k = 20 recent RCTs) — and that spread, together with the substantial I^2, is the statistical fingerprint of exactly the dose and modality inconsistency the Cochrane trajectory already points to. A single pooled number here should be read as a plausible range, not a precise constant.

Two further sources reported standardised mean differences rather than a raw MD and were kept separate rather than force-fitted onto the same scale, but they tell a consistent story. Amjad et al. [8] (JMIR, 2022; PROSPERO-registered, CRD42021256826), pooling 43 trials and 1,893 patients, found SMD
0.45 (95% CI 0.21–0.68, p<.001) for FMA-UE. And the most recent, largest, and arguably most rigorously risk-of-bias-assessed review identified in this search [9] (ScienceDirect, 2026; 34 RCTs, 1,834 patients) found the effect was no longer statistically significant (SMD 0.56, 95% CI −0.02 to 1.15, p=.06; certainty rated “very low”). Taken together with the Cochrane trajectory above, there is a fairly consistent pattern across independent research groups: earlier, smaller pooled estimates tended to favour VR more strongly, and that advantage has narrowed — in the largest, newest reviews, to the point of non-significance — as the evidence base has grown and been assessed more carefully.

3.4 Secondary outcomes

Outcome	Source	Effect estimate (95% CI)	Interpretation
Box and Block Test (dexterity)	Kenea et al. [2]	MD 2.85 (0.70 to 4.99), p=.009	Significant, small effect
Action Research Arm Test	Kenea et al. [2]	MD 3.47 (−0.22 to 7.15), p=.06	Not significant
FMA-UE, subacute stroke	Kenea et al. [2]	SMD 0.92 (0.48 to 1.36), p=.002	Significant, moderate-large
FMA-UE, chronic stroke	Kenea et al. [2]	SMD 0.69 (0.03 to 1.35), p=.03	Significant, moderate
Quality of life	ScienceDirect [9] (34 RCTs)	SMD −0.09 (−0.30 to 0.11), p=.36	Not significant

3.5 Is any of this clinically meaningful?

Statistical significance and clinical significance are not the same thing, and this is really where the whole appraisal lands. The accepted minimal clinically important difference (MCID) for FMA-UE is roughly 7 points — the smallest change a patient or clinician would actually notice as a real improvement in arm function. None of the pooled MD estimates in Figure 2 reached that threshold on their point estimate; only the upper bound of the Liu et al. [7] confidence interval crosses it. In plain terms: VR is very likely doing something measurable, but the average patient getting VR instead of matched-dose conventional therapy is not likely to notice a life-changing difference in their arm because of the technology alone.

Individual landmark trials and study-design features

Beyond the pooled reviews, a handful of individual RCTs recur across almost every meta-analysis in this field and are worth examining directly rather than only through someone else’s pooled number. Table 4 summarises their design and the features most relevant to judging risk of bias — assessor blinding is the main safeguard available in VR trials, since blinding participants and therapists to which arm they are in is rarely possible.

Trial	Design / N	Assessor blinding	Key finding
EVREST pilot (Saposnik et al. [10], Int J Stroke)	Pilot RCT, Wii-based non-immersive VR vs recreational therapy	Yes (outcome assessors blinded)	Established feasibility and safety of low-cost gaming VR; not powered for efficacy

EVREST multicentre (Saposnik et al. [11], Lancet Neurology)	Multicentre, single-blind RCT	Yes	Non-immersive VR produced outcomes comparable to, not better than, recreational therapy
VIRTUES (Brunner et al. [12], protocol 2014)	Multicentre RCT, subacute stroke, dose-matched design	Yes	VR as adjunct to dose-matched conventional therapy; assessed cost-effectiveness and satisfaction alongside motor outcome
Subramanian et al. [13] (Neurorehabil Neural Repair)	RCT, chronic stroke	Yes	VR-based motor learning improved movement quality
Turolla et al. [14]	RCT, intensively-treated patients	Yes	VR showed greater improvement than conventional therapy, but under high-intensity conditions not typical of routine practice

A pattern worth flagging: the trial with the most rigorous, dose-matched design specifically built to isolate the technology's own contribution — VIRTUES — is also the one whose primary purpose was to test cost-effectiveness and satisfaction rather than assume superiority outright, and the EVREST multicentre trial, the largest and best-known of the group, found no superiority over an active comparator. The trials showing the largest advantages for VR (e.g. Turolla [14]) tend to be the ones where the VR arm also received more total therapy time. This is the same dose-confound the Cochrane trajectory in Section 3.2 already surfaces at the review level.

Discussion:

Put together, this synthesis does not overturn the conclusion the original CAT reached narratively — it gives that conclusion numbers. Virtual reality produces a real, measurable improvement in upper-limb motor impairment compared with conventional physiotherapy, somewhere in the region of 3–7 FMA-UE points depending on which review one trusts most, but that range sits right at or below the threshold anyone would call clinically important. The between-review heterogeneity (I^2 = 74.9% among the four pooled estimates) is not noise to be explained away; it is the data telling us that the answer depends heavily on which VR modality, dose, and patient population a given study used, and that a single headline number flattens real clinical variation.

The most useful evidence in this whole review is arguably not any single meta-analysis but the trajectory of the Cochrane review itself across its four updates. As more and larger trials accumulated between 2011 and 2017, the advantage of VR over conventional therapy shrank from a substantial SMD of 0.53 down to a non-significant 0.07 — and then reappeared, at a very similar size to the original effect, the moment VR was used to add extra therapy time on top of usual care rather than replace an equal dose of it. That is about as clean a demonstration as this field is likely to produce that dose, not device, is doing most of the work. The 2025 update, now built on 190 trials, describes its benefits in the same dose-added language. Later, larger reviews identified in this search (Amjad [8] 2022 SMD 0.45; ScienceDirect [9] SMD 0.56, non-significant) sit comfortably within this same story of a real but modest effect that shrinks as the evidence base matures and study quality improves.

For the Indian rehabilitation context that the original CAT was written around, this has a fairly direct practical reading. VR is not a shortcut past skilled therapy, and it should not be marketed or budgeted for as one. Its honest value is as a way to add repetitions and sustain engagement in settings where therapist time is the scarce resource — a home programme supplement, a way to keep a patient practising between supervised sessions, or a tool to make a long session feel shorter. Framed that way, the cost of a VR system is being weighed against additional therapy time gained, not against the

therapist's expertise itself, which is a much easier case to make to a hospital administrator or a family deciding how to spend a limited rehabilitation budget.

Strengths and limitations

The main strength of this review is that every number in it traces back to a real, verifiable, recently published source — nothing here has been estimated or assumed, and the Cochrane trajectory in particular gives the argument a degree of authority that a single new meta-analysis could not. Including the largest and most recent (2026) review also means the synthesis is not resting solely on older, smaller studies that tend to inflate effect sizes.
The main limitation, stated plainly rather than tucked away, is that the quantitative pooling in Section 3.3 works from already-published summary estimates rather than from raw data re-extracted trial by trial — an overview of meta-analyses rather than a de novo, PROSPERO-registered meta-analysis. That is a legitimate design for a synthesis of this scope, but it has less resolution than starting from primary data, and mixing reviews with different inclusion criteria (immersive VR only, versus all VR modalities) is itself a source of the heterogeneity reported above. Publication bias could not be independently assessed across the whole overview; the one source review that tested for it (Kenea et al. [2]) found no significant funnel-plot asymmetry among its own thirteen trials (Egger's p=.08), which is reassuring but does not extend automatically to the other three pooled sources.

Conclusion:

Virtual reality works, in the narrow sense that it produces a statistically real improvement in upper-limb motor function after stroke compared with conventional physiotherapy. It does not work in the sense of being a superior replacement for skilled, task-specific therapy — the evidence, especially the Cochrane review's own trajectory across five updates, points fairly consistently to dose of practice as the real driver of recovery, with VR being one convenient way of delivering more of it. The original CAT's conclusion holds up under quantitative scrutiny: VR is best used as an adjunct that increases repetition, feedback, and engagement within a structured rehabilitation programme, not as a stand-alone substitute for a physiotherapist. For Indian clinical practice specifically, that means VR earns its place in a rehabilitation budget when it demonstrably adds practice time that would not otherwise have happened, and expectations set with patients and families should be calibrated to a modest, incremental benefit rather than a technological breakthrough.

References

1. Nakayama, H.; Jørgensen, H.S.; Raaschou, H.O.; Olsen, T.S. Recovery of Upper Extremity Function in Stroke Patients: The Copenhagen Stroke Study. Arch. Phys. Med. Rehabil. 1994, 75(4), 394–398.
2. Kenea, C.D.; Abessa, T.G.; Lamba, D.; Bonnechère, B. Immersive Virtual Reality in Stroke Rehabilitation: A Systematic Review and Meta-Analysis of Its Efficacy in Upper Limb Recovery. J. Clin. Med. 2025, 14, 1783.
3. Laver, K.E.; George, S.; Thomas, S.; Deutsch, J.E.; Crotty, M. Virtual reality for stroke rehabilitation. Cochrane Database Syst. Rev. 2011, Issue 9, CD008349 (DOI: 10.1002/14651858.CD008349.pub2).
4. Laver, K.E.; George, S.; Thomas, S.; Deutsch, J.E.; Crotty, M. Virtual reality for stroke rehabilitation. Cochrane Database Syst. Rev. 2015, Issue 2, CD008349 (DOI: 10.1002/14651858.CD008349.pub3).
5. Laver, K.E.; Lange, B.; George, S.; Deutsch, J.E.; Saposnik, G.; Crotty, M. Virtual reality for stroke rehabilitation. Cochrane Database Syst. Rev. 2017, Issue 11, CD008349 (DOI: 10.1002/14651858.CD008349.pub4).
6. Laver, K.E.; Lange, B.; George, S.; Deutsch, J.E.; Saposnik, G.; Chapman, M.; Crotty, M. Virtual reality for stroke rehabilitation. Cochrane Database Syst. Rev. 2025, Issue 6, CD008349 (DOI: 10.1002/14651858.CD008349.pub5). This is the current, correct citation — cite this version as the primary Cochrane reference; the 2011/2015/2017 versions above are listed only because they are individually discussed in the historical trajectory in Section 3.2.
7. Liu, et al. Meta-analysis on the impact of virtual reality technology on limb function and quality of life in stroke patients. Signa Vitae 2025.
8. Amjad, I. et al. Effectiveness of Using Virtual Reality-Supported Exercise Therapy for Upper Extremity Motor Rehabilitation in Patients With Stroke: Systematic Review and Meta-analysis of RCTs. JMIR 2022 (PROSPERO CRD42021256826).
9. Effectiveness of virtual reality-based rehabilitation on quality of life and upper extremity functions among stroke survivors: A systematic review and meta-analysis. ScienceDirect, 2026.
10. Saposnik, G.; Mamdani, M.; Bayley, M.; et al. Effectiveness of Virtual Reality Exercises in STroke Rehabilitation (EVREST): pilot RCT. Int. J. Stroke 2010, 5, 47–51.
11. Saposnik, G.; Cohen, L.G.; Mamdani, M.; et al. Efficacy and safety of non-immersive virtual reality exercising in stroke rehabilitation (EVREST): a randomised, multicentre, single-blind, controlled trial. Lancet Neurol. 2016, 15, 1019–1027.
12. Brunner, I. et al. Virtual Reality Training for Upper Extremity in Subacute Stroke (VIRTUES): study protocol.

Neurology 2014/2017.
13. Subramanian, S.K.; Lourenço, C.B.; Chilingaryan, G.; Sveistrup, H.; Levin, M.F. Arm Motor Recovery Using a Virtual Reality Intervention in Chronic Stroke: RCT. Neurorehabil. Neural Repair 2013, 27, 13–23.
14. Turolla, A. et al. Virtual reality for the rehabilitation of the upper limb motor function after stroke, 2013.