METHODOLOGY PAPER

HEALTHCARE RESEARCH · DATA QUALITY

CATI for *Healthcare*

A Methodology for Verifiable, High-Fidelity HCP Data

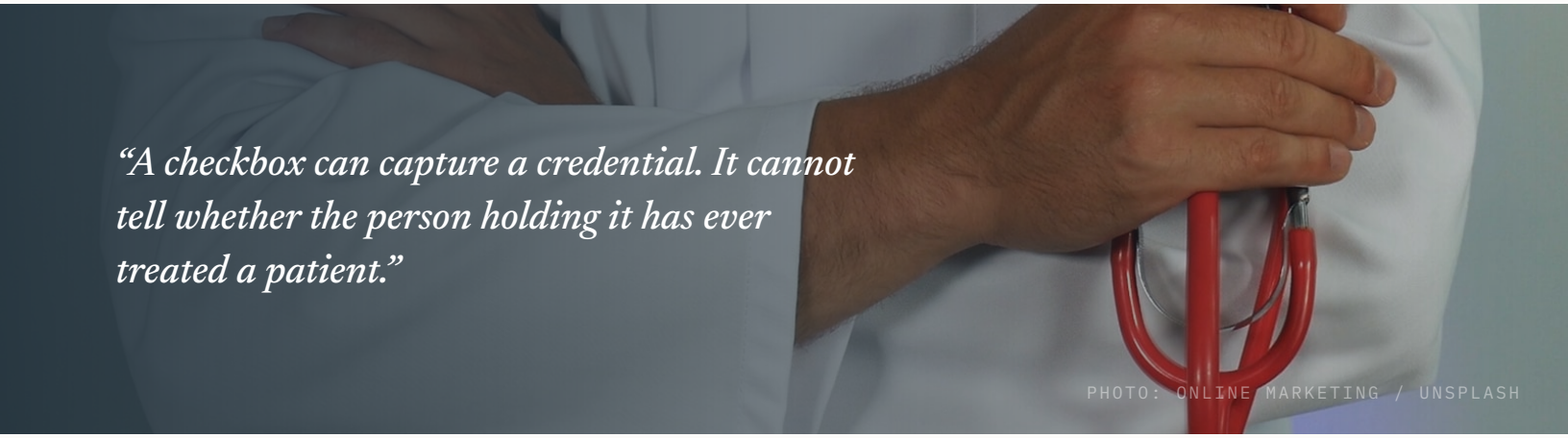
In healthcare research the finding is only as trustworthy as the certainty that a real, practising clinician gave the answer. Self-administered online panels — fast and valuable for many studies — struggle to provide that certainty for small, high-value clinical audiences, where the incentive to fake a credential runs highest. This is the data-quality case for a live, clinically-aware telephone interview: what it verifies, what it captures, and what one fake clinician costs.

PUBLISHED BY
CatalystMR Research Team

SERIES
Methodology Papers

READING TIME
~14 minutes

EDITION
2026



CATALYSTMR
transforming market research data into marketing solutions

ABSTRACT

Healthcare market research carries unusually high stakes: its findings shape launches, formularies, and clinical messaging. Yet it runs on some of the hardest respondents to verify — clinicians whose expertise no checkbox can confirm, drawn from populations small enough that a single fraudulent complete can move a result. Across online research, industry analyses estimate a substantial share of data is discarded over quality and fraud concerns, and peer-reviewed work shows the threat falls **hardest on small target populations** — exactly the rare specialties healthcare studies depend on.

This paper makes the data-quality case for **computer-assisted telephone interviewing (CATI)** in healthcare — not as the only mode, but as the one that delivers **verifiable, high-fidelity** data where it matters most. It maps where HCP panel quality breaks down; shows what a **live clinical interview verifies** that a form cannot; names the **quality a conversation captures**; states honestly **where self-administered modes still win**; and traces **what one fake clinician costs** downstream. It is the healthcare companion to the B2B case of No. 145 and the mode decision of No. 135.

CONTENTS

WHO THIS IS FOR

Healthcare insights leads and brand teams commissioning HCP research; **pharma and medtech market researchers** fielding specialist studies; and **data-quality teams** weighing mode for high-stakes clinical work. No technical background is assumed; practitioners can skip to the benefits in Section 04.

01 THE PREMISE

In healthcare, the data is only as good as the certainty a real clinician gave it.

Most quality frameworks treat fraud as noise to be filtered. In healthcare research it is closer to a single point of failure. The findings steer billion-dollar launches and clinical decisions, yet they rest on respondents whose competence a form cannot confirm and on specialty cells small enough that a few bad completes do not blur the result — they become it. The first question is not "what did clinicians say?" but "were they clinicians at all?"

up to 38%

of online research data discarded over quality and fraud concerns, on industry estimates

KANTAR, VIA RARE PATIENT VOICE

15–30%

commonly cited range for fraudulent responses in online market research

INDUSTRY ESTIMATES

Small n

fraud disproportionately threatens validity in **small** target populations — the rare specialties HCP studies depend on

GORDON ET AL., 2024[1]

These figures describe online research broadly; healthcare concentrates the risk. Rare specialties carry the highest incentive to fake a credential and the smallest base to absorb the damage, and "imposter participants" are a documented, growing threat to online data integrity.[2] None of this makes online panel wrong — for broad, reachable HCP audiences it is fast and valuable. It means that where the audience is small, specialised, and decision-critical, **verification is the product**, and a method that cannot verify is a liability dressed as a dataset.

> THE REFRAMING
>
> Healthcare data quality is not mainly about cleaning answers after the fact. It is about **certainty at the source** — knowing a real, practising clinician is on the other end before a single answer is recorded.

02 THE FAILURE POINTS

Four ways healthcare data fails that general QC misses.

Standard quality controls — deduplication, attention checks, timing — catch generic bad actors. They were not built for the specific ways healthcare data goes wrong, where the threat is not a careless respondent but a convincing impostor in a population too small to hide the damage. Four failure points are distinctly healthcare's.

FAILURE 01

The small-universe fraud magnet

A rare specialty pairs the **highest incentive** to fake a credential with the **smallest base** to absorb it. Fraud concentrates exactly where validity is most fragile.[1]

FAILURE 02

The competence a checkbox can't test

Clinical expertise is not a profile field. A screener captures a **claimed** specialty; it cannot tell whether the respondent can reason, prescribe, or sequence treatment like one who practises.

FAILURE 03

The imposter & professional respondent

Opt-in panels and open links attract serial and **imposter participants** chasing incentives — a documented and increasing threat to online data integrity.[2]

FAILURE 04

The AI-written clinical open-end

Generative tools now produce **plausible-sounding** clinical verbatims that survive a casual read while reflecting no real practice — hollowing out the very answers meant to add depth.

WHY FILTERING ISN'T ENOUGH

Each failure shares a root: the screener and the cleaning pass both act on **what a respondent claims**, not on whether they are who they say. In a small clinical cell, by the time a post-field check flags a problem, the contaminated completes may already be the majority. The durable fix moves verification **upstream**, to the moment of contact.

03 VERIFICATION

What a conversation verifies that a credential alone cannot.

A live interview moves the integrity check from the dataset to the doorway. Within the first minutes of a CATI call, two assurances a screen-and-submit survey can never give are already in hand: the interviewer confirms they are speaking to the named, intended clinician, and a clinically-aware interviewer hears whether that clinician actually practises. Credentials open the door; the conversation confirms the room.

Identity, then competence

Registry matching — NPI, license, association records — verifies that a credential exists; it is the first gate, and the subject of No. 132. A live interview adds the gate a registry can't: **competence in conversation**. An interviewer briefed on the therapeutic area can tell, quickly, whether answers carry the texture of real practice — the right drug names, the plausible sequencing, the caveats a practitioner would raise unprompted.

The tell that a checkbox misses

The risk is not hypothetical. In one widely-cited account, a study's "hemophilia nurses" treated **hand soap as a treatment** — where a real nurse would name factor-replacement therapy. They were not nurses at all, and only a clinical read caught it before it reached a launch decision.

> BUYER'S QUESTION
>
> Ask: **"At what point does someone who actually understands this specialty hear the respondent?"** If the answer is "never," competence was never verified.

04 FIDELITY

What a live clinical interview adds — beyond catching fraud.

Verification is the floor, not the ceiling. Once you know a real clinician is on the line, the conversation captures a fidelity a grid cannot — including benefits buyers rarely price in until they see the difference in the transcript.

01 **Identity certainty at the point of contact** OFTEN OVERLOOKED

You reach and speak with the **named, verified** clinician — not an anonymous profile, a shared login, or a proxy filling in for the incentive.

02 **The real-time clinical-plausibility check** THE DIFFERENTIATOR

A clinically-aware interviewer hears within minutes whether the respondent practises. **Fraud that clears a screener fails a conversation.**

03 **Unaided clinical reasoning & depth** OFTEN OVERLOOKED

Captures how an HCP actually decides — sequencing, rationale, real-world workarounds — in their own words. Richer than a forced grid, and far **harder to fake or AI-generate**.

04 **Engagement through complex instruments**

Interviewer pacing curbs the satisficing and straight-lining that long therapeutic-area batteries invite — so late-survey answers stay as careful as the first.

05 **Specialist reach without fraud-prone recruitment** OFTEN OVERLOOKED

Verified call lists reach rare specialties **without** the open links and outsized incentives that invite imposters in the first place — prevention, not just detection.

05 THE BOUNDARY

Where self-administered modes still win — and why we say so.

A live interviewer is not a free upgrade, and CATI is not the answer to every healthcare study. There is one area where self-administration is genuinely better, and naming it is what makes the rest of this paper trustworthy: a method partner who tells you where their method loses can be believed on where it wins.

Where self-administered wins

- **Sensitive self-report.** Off-label use, adherence failures, and industry influence are disclosed more candidly with no interviewer present — self-administration lowers social-desirability pressure.
- **Scale and speed.** Broad, reachable HCP audiences and large samples are faster and cheaper online.
- **Simple, factual instruments** that need no clarification, pacing, or probing.

So the strongest programs blend

The defensible design is rarely all-or-nothing. Use **CATI where verification, specialist reach, complexity, or reasoning** drive value; use **self-administration where candor, scale, or cost** do — sometimes within one study, under a single master screener. The mode decision is the subject of No. 135; combining modes well is No. 141.

> THE BALANCE TO HOLD
>
> Match the mode to what the study most needs to protect: **certainty and depth**, or **candor and scale**. High-stakes specialist work usually needs the former.

06 THE STAKES

What one fake clinician costs, step by step.

The case for verification is clearest when you follow a single fraudulent complete downstream. In a small specialty cell it does not get diluted — it propagates, and each step magnifies the last until a real decision rests on data that was never real.

STAGE 01 · ENTRY

One fraudulent complete

A non-clinician — or an AI — clears the screener in a small specialty cell and is paid as a qualified respondent.

STAGE 02 · CONTAMINATION

A skewed metric

In a base of dozens, that one complete shifts a therapeutic-area rating, awareness, or preference figure on its own.

STAGE 03 · PROPAGATION

A distorted model

Driver analysis, segmentation, or a forecast inherits the error; the wrong attribute or segment surfaces as the "insight."

STAGE 04 · CONSEQUENCE

A decision made on it

A launch sequence, formulary pitch, or clinical message is shaped by data that was never real — at stakes measured in patients and dollars.

In a general-population study, a handful of fakes wash out in the average. In a **40-physician specialist cell, they are the finding** *– which is why, in healthcare, verification is not housekeeping but the difference between insight and an expensive fiction.*

WHERE THIS SERIES GOES DEEPER

No. 135	**Choosing a Mode for HCP Research.** The full online / CATI / mixed-mode decision this data-quality case informs.
No. 145 · 132	**CATI for B2B & Credible Physician Sample.** The sibling reach-and-judgment case, and the registry credential gate the live interview completes.
No. 138 · 144	**Fraud & straight-lining.** The signals these failures show up as; here the mode itself is the prevention.

CATALYSTMR
transforming market research data into marketing solutions

▲ CONCLUSION

In healthcare, certainty is the product.

Healthcare research asks more of its data than almost any other field, and gives it less margin for error: small populations, high incentives to impersonate, and decisions too consequential to rest on respondents who were never verified. Self-administered online panels remain fast and valuable for broad, reachable audiences and for sensitive self-report, where the absence of an interviewer is an advantage. But where the audience is small, specialised, and decision-critical, a live, clinically-aware telephone interview does what a form cannot — it confirms the clinician is real, hears whether they truly practise, captures the reasoning a grid flattens, and keeps a single impostor from becoming the finding. The strongest programs blend the two, matching mode to what the study most needs to protect. Recognised conduct and service-quality frameworks let buyers ask for that rigour in consistent terms.[3,4]

§ REFERENCES

[1] Gordon, J. H., Fujinaga-Gordon, K., & Sherwin, C. **"Fraudulent Online Survey Respondents May Disproportionately Threaten Validity of Research in Small Target Populations."** Health Expectations 27, no. 3 (2024): e14099. (Fraud concentrates where the population is small.)

[2] Causer, H., et al. **"'Imposter Participants' in Online Qualitative Research, a New and Increasing Threat to Data Integrity?"** Health Expectations 26, no. 3 (2023): 941–944. (Imposter participants as a growing online-integrity threat.)

[3] ICC/ESOMAR. **International Code on Market, Opinion and Social Research and Data Analytics.** ICC & ESOMAR. (Conduct and transparency in fieldwork.)

[4] International Organization for Standardization. **ISO 20252:2019 — Market, opinion and social research, including insights and data analytics.** Geneva: ISO, 2019.

Industry figures. The "up to 38% discarded" and "15–30% fraud" estimates are industry figures attributed to Kantar, reported via Rare Patient Voice (W. Michael, 2025); they describe online research broadly, are not peer-reviewed, and are cited as context, not as CatalystMR or healthcare-specific measurements.

§ ABOUT

CatalystMR

CatalystMR is a global market-research panel and fieldwork partner specialising in **hard-to-reach healthcare, B2B, and niche audiences**. We field **live CATI interviewing with clinically-briefed, monitored interviewers** against verified HCP sample, adding a competence check no checkbox provides — and we advise honestly on when telephone, online panel, or a **blend under one master screener** best protects the result.

Borderline completes are **replaced rather than silently deleted**, so a study keeps both its quality and its planned base.

Compliance posture: aligned to the ESOMAR Code and Guidelines and the ISO 20252 framework; certified under the EU-U.S., UK, and Swiss Data Privacy Frameworks, with personal data siloed from response data.

CATI · HEALTHCARE RESEARCH · HCP DATA QUALITY · PHYSICIAN SAMPLE · ESOMAR CODE · ISO 20252

This paper cites externally-attributed figures (a departure from the series' figure-free norm, at the client's direction); every figure is sourced to a named third party, none is invented, and no operational CatalystMR rate is published. Peer-reviewed sources are cited for the documented concentration of fraud in small populations and the imposter-participant threat; industry figures are labelled as such above. "Aligned to ISO 20252" denotes conformance with the standard's framework, not third-party certification; any turnaround estimate is a modelled feasibility range, not a guarantee.

Tell us your specialty, therapeutic area, and the decision riding on the data — we'll design a verified CATI or blended approach built to earn trust in the result, and return a modelled feasibility range, typically within 24 hours.

sales@catalystmr.com
1-800-819-3130
catalystmr.com