Mass PCR Screening for COVID-19 in Low-Prevalence Populations:

A Critical Reappraisal of the Test-Trace-Isolate Framework Proposed for Japan

George Kujo

Independent Researcher, Tokyo, Japan

ORCID: 0009-0006-6864-0145

Correspondence: george.kujo.irjp@gmail.com

Competing interests: None declared | Funding: None

May 2026 | Version 2.6 (Zenodo Pre-print)

Abstract

Background: During the COVID-19 pandemic, Shibuya and colleagues proposed an aggressive Test-Trace-Isolate (TTI) strategy for Japan, asserting that RT-PCR specificity approached 99.99% and that false-positive concerns were negligible. This study critically examines five fundamental analytical failures of that framework.

Methods: We applied Bayes' theorem using empirically derived Japanese prevalence data (Ministry of Health antibody survey, June 2020: Tokyo seroprevalence 0.1%) and published PCR specificity estimates to calculate the positive predictive value (PPV) across plausible scenarios. We reviewed the published literature on cycle threshold (Ct) values and viral culturing, behavioral responses to COVID-status certification, and documentary evidence from Japan's mass screening operational records, including MHLW quality assurance surveys (FY2020–21).

Results: At 0.1% prevalence and 99% specificity, the PPV collapsed to 6.5%. Even at the claimed 99.99% specificity, the PPV reached

only 87.5%. Japan's Ct=45 threshold detects RNA concentrations approximately one million times lower than those at Ct=25, a level at which infectiousness is epidemiologically negligible. In addition to these previously documented failures, two structural failures were identified. First, negative certificate systems functioned as behavioral disinhibitors: empirical literature documents decreased protective behavior upon receipt of negative certification, and the point-in-time nature of PCR results was not communicated or designed around the time of testing. Second, a pattern of structural implementation failure was identified: the analytical specificity of probe-based RT-PCR was systematically undermined by the program design. Only 18.8% of testing facilities held specialist genetic testing credentials; inter-facility detection sensitivity varied more than fifty-fold; carry-over contamination was documented as a confirmed cause of erroneous results; and volume-based subsidy structures created incentives that were hostile to specimen integrity.

Conclusions: The TTI framework proposed for Japan rested on a conflation of technological specificity with epidemiological utility, and on deployment conditions the implementing system was structurally incapable of guaranteeing. These were foreseeable and analytically identifiable errors. The same research group that dismissed PPV-based governmental caution as 'misinformation' was the source of the more consequential analytical failure. These findings have direct implications for the design of pandemic preparedness policies.

Keywords: COVID-19; PCR screening; positive predictive value; cycle threshold; false positive; test-trace-isolate; Japan; behavioural disinhibition; structural implementation failure; Bayesian inference; pandemic policy

1. Introduction

The COVID-19 pandemic prompted the unprecedented deployment of polymerase chain reaction (PCR) testing as a primary tool for infectious disease control. In Japan, Shibuya and colleagues — including researchers affiliated with King's College London and Japanese academic institutions — proposed a Test-Trace-Isolate

(TTI) strategy modeled on approaches adopted in South Korea and Taiwan, emphasizing the aggressive expansion of PCR testing capacity as the cornerstone of epidemic suppression [1,2].

Central to their argument was the dismissal of false-positive concerns. Shibuya et al. asserted that RT-PCR specificity in infectious disease control contexts approached 99.99% and that the concept of false positives — as applied in clinical diagnostic testing — was fundamentally inapplicable to population-level screening. They further characterized the governmental invocation of false-positive risk as 'misinformation' and 'infodemic' [3].

This paper argues that this position contains five interrelated and consequential analytical failures, each identifiable from publicly available epidemiological and virological data contemporaneous with Shibuya et al.'s publication. These are presented not as retrospective criticism enabled by hindsight but as failures of reasoning that violated established principles of diagnostic test interpretation — principles codified in epidemiology decades before the COVID-19 pandemic [4,5]. Two additional structural failures — the behavioral dynamics of negative certification systems and the systematic gap between analytical and operational specificity in Japan's mass screening programme — are documented for the first time in this version.

This paper further notes, as an institutionally significant fact, that at least some of the analytical errors documented here were contemporaneously identified by members of the public with no specialist training in medicine or epidemiology. This observation invites reflection on the gap between expert consensus and basic logical discipline in scientific discourse during the pandemic.

2. The Positive Predictive Value Problem

A preliminary clarification is needed to address the anticipated objection that Tokyo's 0.1% seroprevalence is irrelevant because TTI targeted a high-prevalence clinical subset rather than the general population. Shibuya and Tokuda et al. explicitly recommended polymerase chain reaction (PCR) testing for

asymptomatic people, including 'asymptomatic people in areas where infections are spreading' and 'travelers from epidemic areas at domestic airports' [1]. They further advocated weekly or biweekly routine testing for essential workers, including teachers, childcare workers, retail workers, and delivery staff [1], and stated that Japan should 'expand testing indications' given its comparatively low testing volume. This is a population-scale screening proposal that far exceeds the high-prevalence clinical subset.

If the proponents later argued that their recommendations applied exclusively to high-risk symptomatic individuals, this would contradict the explicit scope of their published papers. In the populations actually targeted — asymptomatic community members, travellers, and essential workers — the operational prior probability approximates community seroprevalence. The calculations below were calibrated to this population-level prior.

2.1 Theoretical Framework

The relationship between test specificity and epidemiological utility is not fixed; it is a function of the prevalence of the disease in the tested population. This is formalized by Bayes' theorem, from which the positive predictive value (PPV) — the probability that a positive test correctly identifies a truly positive individual — is derived:

PPV = (Sensitivity × Prevalence) / [(Sensitivity × Prevalence) + (1 − Specificity) × (1 − Prevalence)]

This formula reveals the inescapable property that PPV approaches zero as prevalence approaches zero, unless specificity approaches 1.0. This is not a statistical curiosity but a foundational mathematical constraint on the epidemiological utility of diagnostic test applications in low-prevalence populations [4,5,6].

2.2 Application to Japan's Epidemiological Context

The Ministry of Health, Labour and Welfare conducted a population-representative antibody survey in June 2020, confirming a seroprevalence of approximately 0.1% in Tokyo [7]. This is the

most empirically grounded estimate of the true COVID-19 prevalence during the peak TTI advocacy period. This prevalence was applied to the PPV formula:

Specificity	Sensitivity	Prevalence	PPV	False Positive Rate	Note
99.99%	**70%**	**0.1%**	**87.5%**	**12.5%**	**Shibuya et al. claim — best case**
99.00%	**70%**	**0.1%**	**6.5%**	**93.5%**	**Realistic lab conditions [8]**
95.00%	**70%**	**0.1%**	**1.4%**	**98.6%**	**Accounting for inter-facility variation [9]**

Table 1. PPV calculations by specificity scenario at Tokyo, June 2020 seroprevalence (0.1%). Sensitivity is fixed at an illustrative 70% across scenarios, consistent with the 70–80% range stated for RT-PCR by the Japanese Society for Clinical Microbiology [24]; because PPV at low prevalence is governed overwhelmingly by specificity, the conclusions are insensitive to the precise sensitivity value assumed.

Even accepting Shibuya et al.'s most optimistic specificity claim of 99.99%, 12.5% of positive tests would identify non-infectious individuals. Under a realistic specificity estimate of 99%, reflecting inter-facility variation [9], the false-positive rate reached 93.5%.

2.3 The Logical Error in Shibuya et al.'s Dismissal

Shibuya et al. dismissed PPV concerns by reframing the objective: the goal in infectious disease control is population-level epidemic suppression rather than individual diagnosis; therefore, the conventional false-positive/false-negative framework dissolves [2]. This argument is a category error. The epidemiological goal of TTI

does not change the mathematical relationship between specificity, prevalence, and PPV. Whether a positive test is used to manage individual clinical care or to mandate population-level isolation, the probability that it correctly identifies a truly infectious individual remains a function of the prevalence and specificity. Renaming 'false positives' as 'convalescent positives' [2] does not eliminate the social, economic, and iatrogenic harms that accrue to misclassified individuals, as documented in Sections 5 and 6 of this paper.

To engage this position on its own terms, grant its strongest premise outright: let specificity be 100%, so that PPV is also 100%, every positive test marks a genuinely infectious individual, and false positives vanish. The dispute over PPV is then settled in the proponents' favour, and the programme's limitations have still not gone away — they have only moved. At the low prevalence representative of Japan's community screening during most periods (1–3%), the tested population simply contains few infectious people to find. Screen 1,000 people at 2% prevalence with 80% sensitivity and the yield is 16 true positives against 984 negatives, of which four are false negatives who re-enter community circulation while still infectious. Here the binding constraint is not false positives but the small absolute harvest from a low-prevalence pool. Raise prevalence to the 30–60% of a febrile clinic and the false-positive problem recedes, but false negatives grow in absolute number: at 40% prevalence, 1,000 tests at 80% sensitivity yield 320 true positives and 80 false negatives, each holding a negative certificate as they return to households and workplaces. Perfect specificity leaves that last figure untouched [23]. The limitation changes character across the prevalence range; it does not dissolve. Conceding the proponents their most favourable assumption therefore buys no escape from the programme's epidemiological limits — it only alters the form in which those limits appear.

3. The Temporal Validity Problem

A functionally effective TTI strategy requires that positive test results be returned within a timeframe that permits isolation before

the index case's infectious period has substantially elapsed. This requirement was systematically unmet in Japan during the period of maximum TTI.

3.1 The Temporal Arithmetic of COVID-19 Transmission

Parameter	Value	Source
Incubation period (median)	**4–5 days**	**He et al. [10]**
Peak infectiousness	**1–2 days before symptom onset**	**He et al. [10]**
PCR result turnaround, Japan 2020	**2–7 days**	**MHLW reports [7]**
Effective isolation window	**Zero or negative**	**Calculated**

Table 2. Temporal parameters of COVID-19 transmission versus Japan's PCR result turnaround time in 2020.

By the time a positive PCR result was communicated to a Japanese patient in 2020, the window for meaningful pre-isolation transmission had typically closed. The test results were received as a historical document rather than an actionable public health tool. This temporal failure was not a marginal inefficiency but a structural nullification of the central premise of the TTI as an epidemic control strategy. This is directly analogous to issuing a fire alarm after a building has burned down.

This temporal mismatch was identifiable from first principles using publicly available data on incubation periods and result turnaround times. The question was, 'If the result takes longer to return than the infectious period lasts, what is the isolation for?' — was not prominently addressed in Shibuya et al.'s TTI advocacy.

3.2 Structural Absence of Isolation Capacity

A further temporal-structural failure concerns the physical feasibility of isolation. Japan's residential architecture — characterized by compact apartments, multi-generational cohabitation, and limited availability of private individual rooms — meant that 'home isolation' was, in many cases, a policy designation without operational content. An individual instructed to isolate in a shared household cannot prevent transmission to household contacts by mandate. The TTI framework's assumption that positive individuals could achieve effective isolation was contingent on housing conditions that a substantial proportion of the target population did not possess.

3.3 The Paradox of the Negative Certificate: Point-in-Time Testing and Behavioural Disinhibition

The temporal failures documented in Sections 3.1 and 3.2 concern the gap between the test result and isolation: by the time a positive result was communicated, the infectious window had typically closed. This section identifies a structurally distinct but complementary failure that operates on the negative result side of the same testing apparatus.

A PCR test result, whether positive or negative, is a statement about a single point in time. More precisely, it is a statement about the viral RNA present in a specific specimen collected at a specific time. It says nothing about the subject's infection status one hour later, during the journey home from the testing site, or the following morning. This epistemological limitation is not a deficiency of the technology but an inherent property of any point-prevalence diagnostic applied to a dynamic biological process. The University of Tokyo Health Service Center explicitly stated this in its official guidance to travellers: a negative PCR result does not exclude the possibility of SARS-CoV-2 infection.

However, Japan’s mass PCR screening programme was not designed or communicated in a manner consistent with this limitation. The program’s social function was the production of negative certificates — documents that, in practice, served as authorizations for travel, event participation, and commercial

activity. This institutional architecture converts a point-in-time probabilistic statement into a durable warrant for behavior. The certificate did not say: 'this individual was not infected at the moment of specimen collection, under the prevailing conditions of test accuracy and pre-test probability.' It said, functionally: 'this individual is safe to travel.'

The behavioral consequences of this conversion are documented in the empirical literature. A rapid review of 33 studies on COVID-status certification found that protective behaviors (physical distancing, hand hygiene) decreased upon receipt of negative or immunity certificates, regardless of the basis for certification [20]. A study of mandatory surveillance testing programs found that a higher testing frequency was positively and significantly associated with increased participation in risky events, consistent with the risk compensation theory: individuals adjust their behavior to maintain a subjectively acceptable risk level, and a negative test result recalibrates that perception downward [21]. The mechanism is amplified, as the same literature notes, precisely when test inaccuracy is present and results are delayed — the two conditions that characterized Japan's mass screening programme.

This study does not claim that this mechanism is empirically demonstrated for Japan's specific program; the causal relationship between negative certification and increased transmission at the population level remains difficult to isolate from observational data. The claim advanced here is structural and prospective: the TTI framework was designed without accounting for a behavioral dynamic that was theoretically predictable and empirically documented in analogous contexts and operated in the direction of harm.

This structural limitation was quantified by Kucirka et al. (2020), whose prospective cohort study established the relationship between days since infection and false-negative probability in nasopharyngeal PCR specimens [23]. Their data documented false-negative rates of 67–100% during the first three days post-infection, declining to 38% on the day of symptom onset (median day 5), and approximately 20% on day 8 post-infection —

the period of maximum assay sensitivity. These figures are not deficiencies of the testing technology but are consequences of the biological time lag between infection and the accumulation of sufficient viral RNA to cross the detection threshold. No improvement in assay sensitivity or specificity eliminates this time window; even pooled or group-testing schemes designed to mitigate false negatives cannot recover a signal that has not yet accumulated [22]. The implication of a negative certification is direct: an individual infected within 72 hours of specimen collection will receive a negative certificate with a false-negative probability approaching certainty. This individual is not merely undetected; they are actively infectious during the incubation period and will, consistent with the behavioral pattern documented above, act on their certificate's implied assurance of safety — precisely the outcome the programme was designed to prevent.

A further structural irony, identifiable from first principles at the time of program design, concerns the act of obtaining a certificate itself. The behavioral pathway required to receive a negative PCR certificate (travel to a testing site, queuing in waiting areas, and use of public transportation) constitutes a sequence of transmission-exposure events. An individual who was not infected at departure from home might be infected in transit or at the testing facility and would subsequently receive a negative certificate reflecting their pre-exposure status. Multiple cluster events associated with testing facility waiting environments were reported during mass screening in Japan. This transmission pathway was not merely an incidental risk that the program failed to prevent; it was structurally created by the program’s design requirement that individuals physically present themselves at testing locations. The negative certificate produced by this pathway would then function, consistent with this section's central argument, as an authorization for the very mobility that had generated the exposure in the first place.

This dynamic is particularly consequential when considered alongside the PPV analysis in Section 2. An individual in the early incubation window — infectious but not yet generating sufficient viral RNA for reliable detection — would receive a negative

certificate and, consistent with the behavioral pattern documented above, would be more likely, not less, to engage in the mobility that the certificate was ostensibly designed to make safe. The point-in-time limitation and the behavioral disinhibition effect thus operate in the same direction: both increase the transmission risk among certificate holders relative to a counterfactual in which no certification system existed.

The TTI framework proposed by Shibuya and colleagues did not account for this dynamic. The framework assumes that negative certification functions as a transmission-reducing instrument. Behavioral evidence from analogous contexts suggests that it carried a structural potential to function, in part, as the opposite — a potential that the framework neither modelled nor disclosed.

4. Ct Value and Infectivity: The Case of Ct=45

4.1 International Scientific Consensus on Ct Values and Viral Infectivity

Multiple independent studies have established a clear relationship between PCR cycle threshold (Ct) values and viable viral infectivity. The Ct value represents the number of amplification cycles required to produce a detectable fluorescent signal; lower values indicate higher initial viral RNA concentrations.

Bullard et al. (2020) demonstrated that live, culturable virus could not be isolated from specimens with Ct values exceeding 24, regardless of symptom status (OR=0.67 per Ct unit increase; $p<0.001$) [11]. Singanayagam et al. (2020) confirmed that Ct values above 35 yielded successful viral culture in only 8.3% of specimens [12]. Jefferson et al.'s systematic review of viral culture studies confirmed that culturability — and therefore transmission potential — was negligible above Ct=35 [13]. By 2021, an international scientific consensus held that Ct values above 35 represented a practical boundary beyond which the epidemiological risk of transmission was negligible.

4.2 Japan's Operational Ct Value: 45 Cycles

Against this backdrop, Japan's operational standards were strikingly discordant. The National Institute of Infectious Diseases (NIID) protocol adopted Ct=45 as the positivity threshold during the pandemic. This was not merely an advisory guideline; it was embedded in commercially deployed diagnostic software.

The Smart Gene system (Mizuho Medy Co., Ltd.), one of the most widely deployed rapid PCR platforms in Japan, is documented to perform a maximum of 45 amplification cycles, with specimens yielding a signal at any cycle classified as positive [14]. This specification was confirmed in active clinical use as late as August 2024, approximately four years after Bullard et al. established Ct=24 as a practical infectivity threshold, and three years after international convergence on Ct=35.

4.3 The Arithmetic of Ct=45

Comparison	RNA Concentration Difference	Biological Interpretation
Ct=45 vs Ct=40	**2^5 = 32× lower**	**Below any reasonable infectivity threshold**
Ct=45 vs Ct=35	**2^{10} = 1,024× lower**	**International culturability boundary**
Ct=45 vs Ct=25	**2^{20} = ~1,000,000× lower**	**Level of active symptomatic infection**
Ct=45 vs Ct=20	**2^{25} = ~33,554,432× lower**	**Peak viraemia level**

Table 3. RNA concentration differences between Ct=45 (Japan's operational standard) and clinically and epidemiologically relevant Ct reference values.

At Ct=45, the theoretical viral copy number approached or fell below one copy per reaction volume. The biological concept of an 'infectious virus' is definitionally inapplicable at this concentration. A Ct=45-positive individual carries, at most, degraded RNA fragments — the molecular traces of a resolved infection — not replication-competent virions capable of establishing infection in contacts.

4.4 Implications for the Shibuya et al. Framework

Shibuya et al. put RT-PCR specificity in infectious disease control at 99.99% [2], but never stated the Ct value at which that figure was measured or assumed. The omission is not a technicality; it is where the argument turns.

Specificity is not a fixed property of PCR technology but a function of the cycle threshold applied. As the Ct value increases, the analytical sensitivity increases, whereas the effective specificity for identifying infectious individuals decreases monotonically. At Ct=45

— Japan's documented operational standard — the question to be asked is not whether specificity reaches 99.99% for the presence of any RNA, but whether the assay retains meaningful specificity for identifying individuals capable of transmitting SARS-CoV-2. Four years of accumulated international evidence clearly answer this question in the negative.

The combination of low population prevalence (Section 2) and Ct=45 (this section) produces a compounded failure. The PPV for identifying truly infectious individuals approaches zero by any realistic calculation. This is not a theoretical edge case but a description of Japan's operational reality that persisted through the pandemic and beyond.

5. The Iatrogenic Risk of Quarantine Facilities

The TTI framework assumes that the isolation of positive individuals prevents onward transmission. This assumption holds only if isolated individuals are infectious and if isolation successfully prevents contact between positive and negative individuals. Both assumptions were violated in Japan's implementation.

5.1 The Paradox of Quarantine-Acquired Infection

When false-positive individuals — identified as positive due to PPV collapse at low prevalence (Section 2) or due to Ct=45 detection of non-infectious RNA fragments (Section 4) — are placed in quarantine facilities alongside genuinely infectious individuals, non-infected persons are exposed to active viral transmission. The intervention designed to prevent infection becomes a mechanism of infection delivery.

This is not a theoretical issue. Japan's hotel quarantine system (Hotel Therapy), implemented from 2020 onward, co-located all PCR-positive individuals, irrespective of infectiousness, symptom status, or Ct value. At prevalence levels consistent with the calculations in Section 2, a substantial proportion of hotel quarantine residents were non-infectious at admission, creating

conditions for facility-acquired infection that would not have occurred had those individuals remained at home.

5.2 Parallel Evidence from Healthcare Settings

Similar dynamics have been extensively documented in hospitals during the same period. Studies from the United Kingdom estimated that 20–25% of COVID-19 cases in hospitalized patients were nosocomially acquired [15]. The mechanism in quarantine hotels is directly parallel: congregating individuals with heterogeneous infectious statuses in a shared enclosed environment generates transmission opportunities that geographically dispersed home isolation would not.

The Shibuya et al. The TTI framework, which explicitly advocated for expanded quarantine facility capacity [1], made no provision for this iatrogenic risk. The risk was foreseeable based on the basic principles of infection control and contemporaneous hospital epidemiology data.

6. Opportunity Cost: The Crowding-Out of General Healthcare

The framework's costs did not stop at the false-positive individuals it harmed directly. By hospitalising or quarantining every PCR-positive person regardless of clinical severity, infectiousness, or Ct value, the policy drew finite capacity away from the rest of Japan's healthcare system.

Domain	Documented Impact	Source
Cancer screening	**10–30% reduction in participation, FY2020**	**Machii & Takahashi [16]**
Acute MI management	**Increased door-to-balloon time; elevated in-hospital mortality**	**Japanese Circulation Society [17]**

Home deaths during isolation	**Multiple documented cases, 2021 Delta wave**	**MHLW records [7]**

Table 4. Documented opportunity costs of the COVID-19 PCR policy on general healthcare in Japan.

The hospitalization of individuals testing positive at Ct values of 40–45 — individuals with, at most, trace non-replicating viral RNA — occupied beds unavailable for patients with acute myocardial infarction, stroke, trauma, and malignancy requiring timely intervention. Cancer screening participation decreased by 10–30% in FY2020 [16]. These are not hypothetical harms; they are documented in the peer-reviewed literature and government records.

This is a textbook instance of opportunity cost in the allocation of healthcare resources. It was foreseeable from first principles at the time TTI policy was being formulated, and it was in fact identified contemporaneously by non-specialist observers who noted: 'If you hospitalise people who are PCR positive but not sick, you delay treatment for people who are sick.' The Shibuya et al. framework does not systematically address this trade-off.

7. The Implementation Gap: Structural Divergence between Analytical Specificity and Operational Reality

Arguments 1–4 demonstrate the failures in the theoretical foundations of the TTI framework. This section documents a further and independent failure: the systematic divergence between the technical conditions under which probe-based RT-PCR achieves its claimed performance and the conditions that were actually obtained in Japan's mass PCR screening programme.

7.1 Framing: Two Concepts of Specificity

A technically sophisticated defence of mass PCR screening contends that probe-based real-time RT-PCR exhibits near-perfect analytical specificity by virtue of its three-oligonucleotide recognition mechanism: forward primer, reverse primer, and fluorescent probe acting in concert. Under idealized laboratory conditions, the probability of a false positive arising from random sequence complementarity is astronomically small, and this property has been invoked to argue that Bayesian PPV calculations premised on imperfect specificity are conceptually misapplied to the technology.

This argument, while technically accurate within its own domain, conflates two distinct constructs: analytical specificity, which refers to the performance of the assay with purified target sequences under controlled conditions, and clinical specificity, which refers to the observed false positive rate in real-world deployment across heterogeneous specimens, facilities, and operators. The former is a chemical property, whereas the latter is a property of the system. Mass screening programs are systems, not assays.

The empirical evidence supports this distinction. External quality assessment studies have documented false-positive rates ranging from 0% to 16.7% across real-world laboratories, with an interquartile range of 0.8–4.0%. The US Food and Drug Administration does not require 100% specificity for PCR-based screening authorization; a threshold of ≥98% is accepted, implicitly acknowledging operational variance. Post-market surveillance has identified individual commercial kits with a specificity as low as 81.8%. These figures are not theoretical; they represent the documented performance outcomes of deployed systems.

7.2 Structural Depression of Pre-test Probability by Programme Design

The analytical specificity argument, even if accepted on its own terms, cannot resolve a prior and independent problem: the systematic depression of pre-test probability by the design of Japan's mass screening program itself.

Japan's publicly subsidized free PCR testing programme, as formally defined by the Tokyo Metropolitan Government and adopted by implementing entities nationwide, targeted asymptomatic individuals engaged in economic activity — specifically, workers in the service sector seeking to obtain certification of negative status to participate in commercial transactions. This population-level design decision had a direct and predictable consequence for test performance: it maximized the proportion of true negatives in the tested population, thereby minimizing the positive predictive value regardless of the assay-level specificity.

This is not a critique of this technology. This is a critique of the epidemiological logic governing its deployment. A test with 99.9% analytical specificity applied to a population with 0.1% true prevalence produces a positive predictive value of approximately 50% — a result indistinguishable from a coin toss. The program's design guaranteed a low pre-test probability as a structural feature, not as an incidental outcome.

7.3 The Non-medical Character of Operational Entities

The theoretical performance of probe-based RT-PCR is predicated on its implementation by trained molecular biologists operating under laboratory-quality management systems. Japan's mass PCR screening program was not implemented under these conditions.

Documentary evidence from the program's operational period establishes that PCR testing sites were operated by entities whose primary business activities were unrelated to medical diagnostics. Publicly available corporate materials from this period identified estate clearance and house-cleaning companies as entities operating certified testing sites. The specimen collection and processing workflow at these facilities was conducted by staff without documented molecular biology training, a circumstance confirmed by consumer complaint records reported in investigative journalism. A vending machine-based PCR testing system, in which saliva specimens were collected by consumers and returned by

post to a non-medical retailer for processing, was documented in local press reporting.

The MHLW's commissioned quality control survey for FY2020–21 found that only 18.8% of facilities conducting PCR-based testing held specialist credentials for genetic testing. This figure implies that the overwhelming majority of mass testing capacity was operated by entities that did not meet the institutional prerequisites assumed by the analytical specificity argument.

7.4 Volume Incentive Structures and the Collapse of Specimen Integrity

The subsidy structure governing Japan's mass PCR screening program created financial incentives misaligned with diagnostic accuracy. Under the per-test reimbursement model, operating entities receive a fixed government subsidy for each test processed. This structure rewarded volume maximization and was indifferent to the result quality.

Investigative journalism has documented the downstream consequences of this incentive architecture. Broker networks emerged that solicited specimen fabrication at a price of approximately ¥3,000 per test, with operatives instructed to submit identification information without collecting actual specimens. Whether the resulting processed 'results' entered official COVID-19 surveillance databases is not fully established, but the documented existence of such networks demonstrates that the specimen integrity assumptions underlying analytical specificity calculations were structurally compromised at the program level.

Even in the absence of deliberate fabrication, the incentive to maximize throughput created pressure against the time-intensive pre-analytical protocols — specimen labelling verification, storage temperature management, and transport chain documentation — that quality PCR results require. The per-test subsidy model did not compensate for these pre-analytical investments, creating structural incentives to abbreviate them.

7.5 Pre-analytical Contamination Risk

The MHLW's FY2021 quality assurance survey for clinical testing facilities documented carry-over contamination as a confirmed cause of erroneous PCR results in two facilities and two cases. Carry-over contamination, the transfer of amplified products from a previous positive reaction into a subsequent test run, is a well-characterized failure mode in PCR diagnostics. This is not a failure of the probe's recognition chemistry but a failure of the preanalytical and postanalytical physical environments.

The survey additionally found that facility-level accuracy for COVID-19 PCR testing ranged from 83.3% to 100% in the lowest performance tier, and that inter-facility variation in detection sensitivity spanned more than fifty-fold. These figures were derived from facilities operating within the regulated clinical testing sector in Japan. The performance variance within the regulated sector suggests that the variance in the unregulated mass screening sector was likely considerably greater.

7.6 Structural Implementation Failure as an Analytical Category

The evidence presented in Sections 7.2–7.5 converges on a concept we term structural implementation failure: a condition in which the theoretical justification for a policy intervention depends on technical prerequisites that the implementing system is structurally incapable of guaranteeing.

Structural implementation failure differs from ordinary implementation shortcoming. The latter implies that a policy could have worked if executed better, while the former implies that the gap between ideal and actual conditions was not a contingent failure of execution but a necessary consequence of the policy's institutional and incentive architecture. Japan's mass PCR screening programme exhibited structural implementation failure in this sense. The analytical specificity argument is conditioned on 'appropriately designed and managed' operational environments. The program was designed in ways that systematically prevented

those conditions from being met: by targeting asymptomatic populations with structurally low pre-test probability; by delegating implementation to non-medical entities without specialist credentials; by creating per-test volume incentives hostile to pre-analytical quality; and by operating without the quality management infrastructure that clinical RT-PCR requires.

This argument is independent of Arguments 1–4. It does not require contested assumptions regarding specificity values, Ct thresholds, or comparator interventions. It requires only that the conditions necessary for the technology to perform as its proponents claimed were absent, and that their absence was structurally guaranteed by the program's design. This requirement was satisfied.

8. Discussion

8.1 The Compound Nature of the Failure

The five analytical failures documented in this study — PPV collapse at low prevalence, temporal invalidity of TTI, Ct-infectivity dissociation at Ct=45, opportunity cost, and structural implementation failure — are not independent errors. They are structurally interconnected, with each compounding the harm generated by the others. A policy that generates large numbers of false positives (Section 2), sends those false positives to quarantine facilities too late to interrupt transmission (Section 3), co-locates them with genuinely infectious patients (Section 5), simultaneously reduces the healthcare capacity available for the non-COVID sick (Section 6), and was implemented through a system structurally incapable of delivering the analytical performance its justification required (Section 7) is not merely suboptimal. In the most precise clinical sense, it is iatrogenic.

The Ct=45 threshold (Section 4) was the linchpin. Had Japan adopted the international consensus threshold of Ct=35, the false-positive problem would have been substantially mitigated, quarantine bed pressure would have been reduced, and iatrogenic co-exposure risk would have been diminished. The persistence of

Ct = 45 as Japan's operational standard through at least August 2024, long after the relevant scientific literature had accumulated, represents an institutional inertia that transcends the specific errors of the Shibuya et al. framework.

8.2 The Rationale for Ct=45 and the Limits of the Counter-Argument

Rigorous analysis requires engagement with the strongest version of the opposing argument. Proponents of Ct=45 and broad TTI could reasonably invoke three justifications: precautionary posture under genuine early uncertainty, political demand for zero missed cases, and the fundamental sensitivity-specificity trade-off. However, these arguments do not survive contact with the timeline. By mid-2020, peer-reviewed evidence that Ct values above 25-30 were associated with non-culturable, non-transmissible viruses had begun to accumulate [11,12,13]. By 2021, an international consensus was forming around Ct thresholds well below 40. However, Japan's operational standard of Ct=45 persisted through at least August 2024. The failure under examination is not the initial adoption of a precautionary threshold amid genuine uncertainty. This is institutional inertia in updating the threshold as the evidence base matured.

8.3 The Infodemic Inversion

A particularly striking aspect of this case is the rhetorical posture adopted by the Shibuya group toward critics of mass PCR testing. Shimizu and colleagues published a paper characterising the Japanese government's invocation of false positive risk as 'misinformation' and contributing to an 'infodemic' [3]. However, as demonstrated in Sections 2 through 5, the government's concern was analytically grounded. At Tokyo's June 2020 seroprevalence of 0.1%, a specificity of 99% yields a PPV of 6.5%, meaning that 93.5% of positive tests would identify non-infectious individuals. This is not a policy opinion; it is arithmetic derived from Bayes' theorem, from equations codified in epidemiology since Vecchio (1966) [4]. The study that accused others of producing an infodemic

did not contain a single Bayesian calculation. The irony is analytically precise: the group that characterized PPV-based caution as misinformation was itself the source of a more consequential analytical error.

8.4 The Role of the Non-Specialist Observer

The reasoning required to identify the failures documented in this study is not abstruse. It requires only asking, What is this test for? (purpose) When will the result arrive? (timeliness) What does a positive result mean in terms of population prevalence? (PPV) What is the cost of acting on this result? (opportunity cost) Does the implementing system actually meet the conditions required by the technology? (implementation fidelity) These are questions of basic logical structure that require no specialist training to formulate, although they do require institutional incentives that reward their honest pursuit.

The gap between lay common sense and expert policy recommendations in this case illustrates not the superiority of lay judgment but the degree to which institutional incentives, professional specialization, and the social dynamics of pandemic-era discourse displaced basic analytical discipline. No single expert community was able to identify the compound failure. This is an argument for interdisciplinary pandemic preparedness frameworks, not for the primacy of non-expert judgment over specialist expertise.

8.5 Limitations

This study is a critical analytical review, not an independent empirical study. The prevalence estimates used in Section 2 were derived from antibody surveys with methodological limitations, including potential underascertainment and temporal variation. Sensitivity and specificity estimates were drawn from published studies conducted in various clinical contexts. The Ct=45 documentation in Section 4 relies on peer-reviewed literature and contemporaneous clinical records rather than a systematic survey

of all Japanese testing facilities. The implementation failure documentation in Section 7 relies on publicly available investigative journalism, corporate materials and government surveys. The behavioural disinhibition evidence discussed in Section 3.3 comes from studies conducted outside Japan; no study has directly demonstrated the mechanism within Japan's own mass screening programme, and the argument there is accordingly structural and prospective rather than empirically established for this case. The analysis also does not model the counterfactual of no TTI implementation, and so does not attempt to quantify net harm relative to a world in which the policy had not been adopted. These limitations do not undermine the core analytical arguments, which are mathematical in nature and hold across the plausible ranges of the input parameters.

8.6 Cross-Domain Convergence: Institutional Inertia as a Structural Pattern

The five analytical failures documented in this paper share a common structural feature: each was identifiable from the evidence and reasoning available at the time of policy formulation, yet the policy persisted. A companion study by the author identifies the same structural dynamics in two further domains of Japanese COVID-19 policy. In obstetric guideline design, Kujo [18] demonstrated through Emergency Guideline Elasticity Theory (EGET) that high-elasticity guideline language enabling non-evidence-based caesarean sections remained unrevised through Waves 2–4. The Crisis Policy Drift mechanism identified in that study describes the Ct=45 persistence documented here with equal precision. In the domain of hospital visitation policy, Kujo [19] formalizes Prolonged Precaution Drift (PPD) as the continuation of categorical precautionary measures beyond the reversal point at which marginal infection-control benefit no longer justifies accumulated harm. Taken together, these three analyses provide convergent evidence that pandemic institutional inertia operated as a cross-domain structural pattern in Japan.

9. Conclusions

At its root, the Test-Trace-Isolate framework Shibuya and colleagues proposed for Japan treated the specificity of a technology as if it were the same thing as epidemiological usefulness. Declaring RT-PCR specificity of 99.99% sufficient to dismiss false-positive concerns took that 99.99% to hold independently of how common the disease actually was — which is precisely what the mathematics of diagnostic testing forbids.

That foundational error did not travel alone. The framework also said nothing about how Ct values relate to infectivity, about the point-in-time epistemology of negative certificates and the behaviour they license, or about the distance between the laboratory conditions in which analytical specificity is measured and the conditions under which Japan actually ran its mass screening. Japan held Ct=45 as its operational threshold — written into national guidelines and into commercial hardware, and still in clinical use as late as August 2024 — a cutoff that registers viral RNA at roughly one-millionth the concentration of active infection. Low population prevalence and a Ct of 45 together drove the PPV for identifying genuinely infectious individuals down to a figure no realistic calculation can rescue.

The consequences were not merely theoretical. False-positive individuals were quarantined alongside genuinely infectious patients, creating conditions for facility-acquired infections. Hospital beds were occupied by PCR-positive individuals with non-replicating viral RNA, while patients with acute cardiac and oncological conditions faced delayed care. Cancer screening rates fell by 10–30% (FY2020). The implementing system was structurally incapable of guaranteeing the analytical conditions required by the framework’s justification.

These were foreseeable harms, analytically identifiable using the data and frameworks available at the time policy recommendations were made. Their non-identification — and active suppression through accusations of misinformation — by the research community that designed and advocated for Japan's TTI framework

represents institutional inertia with direct implications for pandemic preparedness.

The central lesson is that PCR testing is not without value. However, any diagnostic technology, however analytically precise, can generate net harm when deployed without accounting for the mathematical relationship between prevalence and predictive value, the temporal structure of the disease being targeted, the Ct threshold applied, the epistemological limits of negative certification, the co-location risks of isolation facilities, the opportunity costs imposed on the broader healthcare system, and the operational conditions of the implementing system. Pandemic preparedness frameworks that ignore these interactions do so at the expense of the population.

References

[1] Shimizu K, Kondo T, Tokuda Y, Shibuya K. An open letter to Japan's new Prime Minister. Lancet. 2020;396(10259):e57. doi:10.1016/S0140-6736(20)32033-X. PMID:33002428.

[2] Tokuda Y, Shibuya K, Oguro K. Priority of SARS-CoV-2 test, trace, and isolation in Japan. J Gen Fam Med. 2021;22(1):1–2. doi:10.1002/jgf2.409. PMID:33362983.

[3] Shimizu K, Shibuya K, Tokuda Y. COVID-19 infodemic about nucleic acid amplification tests in Japan. J Gen Fam Med. 2022;23(1):1–2. doi:10.1002/jgf2.504. PMID:34904060.

[4] Vecchio TJ. Predictive value of a single diagnostic test in unselected populations. N Engl J Med. 1966;274(21):1171–1173. doi:10.1056/NEJM196605262742104. PMID:5934954.

[5] Meehl PE, Rosen A. Antecedent probability and the efficiency of psychometric signs, patterns, or cutting scores. Psychol Bull. 1955;52(3):194–216. doi:10.1037/h0048070. PMID:14371890.

[6] Watson J, Whiting PF, Brush JE. Interpreting covid-19 test result. BMJ. 2020;369:m1808. doi:10.1136/bmj.m1808. PMID:32398230.

[7] Ministry of Health, Labour, and Welfare, Japan. Kotai hoyu chosa no kekka ni tsuite [Results of the SARS-CoV-2 antibody seroprevalence survey]. Tokyo: MHLW; published June 16, 2020. A survey was conducted in the first week of June 2020 on approximately 3,000 randomly sampled residents each in Tokyo, Osaka, and Miyagi prefectures (Tokyo seroprevalence 0.10%). https://www.mhlw.go.jp/stf/newpage_11892.html (accessed 20 May 2026). Neutralisation test results: https://www.mhlw.go.jp/stf/newpage_12409.html [In Japanese]

[8] Dinnes J, Deeks JJ, Adriano A, et al. Rapid, point-of-care antigen, and molecular-based tests for the diagnosis of SARS-CoV-2 infection. Cochrane Database Syst Rev. 2020;8(8):CD013705. doi:10.1002/14651858.CD013705. PMID:32845525.

[9] Ministry of Health, Labour and Welfare, Japan (Commissioned project). Shingata korona uirusu kansensho no PCR kensa to ni kakaru seido kanri chosa gyomu hokokusho [Report on the quality control survey of COVID-19 PCR and related testing]. Tokyo: MHLW; April 2021. Of the 563 facilities surveyed, 18.8% [106/563] held specialist genetic testing credentials. https://www.mhlw.go.jp/content/000769978.pdf (accessed 20 May 2026). [In Japanese]

[10] He X, Lau EHY, Wu P, et al. Temporal dynamics of viral shedding and transmissibility of COVID-19. Nat Med. 2020;26(5):672–675. doi:10.1038/s41591-020-0869-5. PMID:32296168.

[11] Bullard J, Dust K, Funk D, et al. Predicting infectious SARS-CoV-2 from diagnostic samples. Clin Infect Dis. 2020;71(10):2663–2666. doi:10.1093/cid/ciaa638. PMID:32442256.

[12] Singanayagam A, Patel M, Charlett A, et al. Duration of infectiousness and correlation with RT-PCR cycle threshold values in cases of COVID-19 in England. Euro Surveill. 2020;25(32):2001483. doi:10.2807/1560-7917.ES.2020.25.32.2001483. PMID:32794447.

[13] Jefferson T, Spencer EA, Brassey J, Heneghan C. Viral Cultures for Coronavirus Disease 2019 Infectivity Assessment: A Systematic Review. Clin Infect Dis. 2022;73(11):e3884–e3899. doi:10.1093/cid/ciaa1764. PMID:33270107.

[14] Kurihara Y, Kiyasu Y, Akashi Y, et al. Evaluation of a novel digital immunochromatographic assay with silver amplification to detect SARS-CoV-2. J Infect Chemother. 2021;27(10):1493–1497. doi:10.1016/j.jiac.2021.07.006. PMID:34294528. [Note: reference PCR performed by Mizuho Medy using 45 amplification cycles per NIID Manual Ver.2.9.1]

[15] Read JM, Green CA, Harrison EM, et al.ISARIC4C investigators. Hospital-acquired SARS-CoV-2 infection during the first COVID-19 pandemic wave in the UKthe Lancet. 2021;398(10305):1037-1038. doi:10.1016/S0140-6736(21)01786-4. PMID:34391505.

[16] Machii R, Takahashi H. Japanese cancer screening programs during the COVID-19 pandemic: Changes in participation between 2017-2020. Cancer Epidemiol. 2023;82:102313. doi:10.1016/j.canep.2022.102313. PMID:36508964.

[17] Watanabe Y, Miyachi H, Mozawa K, et al. Impact of the COVID-19 Pandemic on ST-elevation Myocardial Infarction: A Single-center Experience in Tokyo. Intern Med. 2021;60(23):3693-3700. doi:10.2169/internalmedicine.8220-21. PMID:34565777.

[18] Kujo G. Language structure of cesarean section indications and its institutional expansion during the COVID-19 pandemic: Emergency guideline elasticity theory (EGET). Zenodo. 2026. https://doi.org/10.5281/zenodo.18742139

[19] Kujo G. Visitor restrictions beyond the emergency phase: lessons from Japan for proportionality and sunset mechanisms. J Med Ethics. 2026. doi:10.1136/jme-2026-111889

[20] Drury J, Mao G, John A, et al. Behavioral responses to Covid-19 health certification: a rapid review. BMC Public Health. 2021;21:1205. doi:10.1186/s12889-021-11166-0. PMID:34162364.

[21] Jones Ritten C, Thunström L, Cherry T, Wulfhorst JD. Frequent mandatory COVID-19 testing may have increased risky behavior. PNAS Nexus. 2022;1(5):pgac247. doi:10.1093/pnasnexus/pgac247. PMID:36712346.

[22] Alizad-Rahvar AR, Vafadar S, Totonchi M, Sadeghi M. False Negative Mitigation in Group Testing for COVID-19 Screening. Front Med (Lausanne). 2021;8:661277. doi:10.3389/fmed.2021.661277. PMID:34095171.

[23] Kucirka LM, Lauer SA, Laeyendecker O, Boon D, Lessler J. Variation in False-Negative Rate of Reverse Transcriptase Polymerase Chain Reaction–Based SARS-CoV-2 Tests by Time Since Exposure. Ann Intern Med. 2020;173(4):262–268. doi:10.7326/M20-1495. PMID:32422057.

[24] Japanese Society for Clinical Microbiology. Shingata korona uirusu (SARS-CoV-2) no kensaho ni kansuru joho hasshin [Information on testing methods for SARS-CoV-2]. Tokyo: JSCM. https://www.jscm.org/uploads/files/guideline/291.pdf (accessed 27 May 2026). [In Japanese; states RT-PCR sensitivity of 70–80%]

Disclosure Statement

Competing Interests

None declared.

Funding

None.

AI Assistance Declaration

The preparation of this manuscript involved the use of Claude (Anthropic), a large language model AI assistant, for literature organization, mathematical calculation verification, manuscript

drafting, and document formatting. All analytical frameworks, core arguments, and intellectual content originated from the human author. AI was used as a research and writing tool. The author takes full responsibility for the accuracy and integrity of all the content.

Revision History

Version	Date	Changes
v1.0	**January 2026**	**Initial Zenodo preprint. Core arguments: PPV collapse (Section 2), temporal invalidity (Section 3), Ct=45 infectivity dissociation (Section 4).**
v2.0	**January 2026**	**Added Section 5 (Iatrogenic Risk of Quarantine Facilities) and Section 6 (Opportunity Cost) to the manuscript. The discussion was expanded with Infodemic Inversion analysis (Section 7.3).**
v2.1	**February 2026**	**Section 7.4 (Non-Specialist Observer) has been added. Strengthened counterargument engagement in Section 7.2.**
v2.2	**February 2026**	**Section 7.6 (Cross-Domain Convergence) has been added. Integration with**

		companion papers on EGET and PPD.
v2.3	**February 2026**	**A technical clarification note was added. Bilingual (Japanese/English) abstracts produced for Zenodo upload. Minor typographic corrections have been made.**
v2.4	**March 2026**	**Two new sections were added: Section 3.3 (Paradox of the Negative Certificate — point-in-time epistemological limitation and behavioral disinhibition effect, with Kucirka et al. 'sfalse negative rate data [23] and testing-site transmission pathway); Section 7 (The Implementation Gap — structural divergence between analytical specificity and operational reality, including 7.1–7.6). Section 2.3 was expanded with a specificity-100% concession argument demonstrating that structural limits persist at both low and high prevalence regardless of false positive elimination. The**

		introduction has been updated to reflect the five analytical failures. Abstract revised accordingly. New references [20]–[23] have been added. Section 3.2 (Structural Absence of Isolation Capacity) has been added. The Discussion has been renumbered as Section 8, and the Conclusions has been renumbered as Section 9.
v2.6	May 2026	Organized citations and references: Added Reference [24] (Japanese Society for Clinical Microbiology) as the source for the illustrative assumption of 70% sensitivity in Table 1. Integrated previously uncited Reference [22] into Section 3.3. Other corrections: Standardized the cancer screening figures in the conclusion to 10–30% (FY2020) to ensure consistency with Section 6 and Table 4. Converted the exponents in Table 3 to native superscript formatting to prevent character corruption. Revised the wording in Sections 2.3, 4.4, 6, and

		9 for style and clarity. Added two qualifications to the limitations section: that the evidence on behavioral disinhibition derives from studies outside Japan, and that the counterfactual scenario of non-implementation was not modeled. Removed the unsourced row on “availability of general hospital beds” from Table 4.