International Journal of Law, Policy and Social Review
www.lawjournals.net
Online ISSN: 2664-6838
Received: 23-08-2026, Accepted: 22-09-2026, Published: 07-10-2026
Volume 8, Issue 4, 2026, Page No. 4-8

Measuring justice without controlling it: Performance review, data governance and tribunal independence under the Tribunals Reforms Act, 2026

Prabhash Kumar Shukla, Mahi Tiwari
National Law Institute University, Bhopal, Madhya Pradesh, India

Abstract
The Tribunals Reforms Act, 2026 creates a National Tribunals Commission (NTC) and gives it a seemingly benign but constitutionally significant function: to review the performance of tribunals and maintain a National Tribunals Data Grid. This paper argues that the new performance-review power should not be treated as ordinary administrative supervision. Tribunals exercise judicial power, and an efficiency regime that rewards disposal numbers, reversal rates or outcome patterns can indirectly influence adjudication as effectively as direct executive control. Drawing on the Supreme Court's tribunal-independence jurisprudence, the 2026 statutory framework, Indian studies on judicial efficiency, and comparative court-performance systems, the paper develops a constitutional distinction between institution-level administration metrics and adjudicator-level decisional metrics. It proposes a three-tier 'independence-compatible performance framework': green metrics that may be published and acted upon; amber metrics that require contextual safeguards; and red metrics that should not be tied to appointment, reappointment, discipline or budgetary consequences. The central claim is that the NTC can become a genuine institutional safeguard only if performance review is designed to measure the conditions of justice rather than the substantive exercise of judicial judgment.

Keywords: Tribunals, judicial independence, National Tribunals Commission, performance measurement, data governance, administrative law, Tribunals Reforms Act 2026

Introduction

India's tribunal system has repeatedly generated the same constitutional anxiety: bodies created to exercise judicial or quasi-judicial power have often remained administratively dependent on the executive departments whose decisions they review. The Supreme Court's tribunal jurisprudence—from S.P. Sampath Kumar and L. Chandra Kumar to the successive Madras Bar Association decisions—has therefore focused on appointments, tenure, qualifications, infrastructure and financial autonomy. In November 2025, the Supreme Court again held that executive dependence threatened tribunal independence and directed the creation of a National Tribunals Commission as an institutional solution [1].

Parliament responded with the Tribunals Reforms Act, 2026. Section 4 assigns the NTC four principal functions: conducting selection processes for covered tribunals, reviewing tribunal performance and preparing an annual report, overseeing inquiries into complaints, and developing and maintaining a National Tribunals Data Grid [2]. The statutory move from fragmented parent-ministry administration to a central commission is significant. Yet the phrase 'review the performance of the Tribunals' is undefined. That silence matters because performance measurement is never neutral: the choice of metric can change institutional behaviour.

This paper addresses a question that has received less attention than appointments and tenure: how can the NTC measure tribunal performance without creating a new mechanism of decisional control? The paper's thesis is that the Constitution permits robust measurement of institutional capacity and procedural efficiency, but sharply constrains the use of metrics that evaluate adjudicative outcomes or attach career consequences to numerical productivity. The distinction should be embedded in rules made under the 2026 Act and in the design of the National Tribunals Data Grid.

Research Question, Method and Contribution

The principal research question is: what kinds of performance indicators may the NTC constitutionally collect, publish and use while preserving tribunal independence? The paper adopts a doctrinal and comparative methodology. It first derives constitutional constraints from Indian tribunal jurisprudence; second, it identifies the administrative problems that justify measurement; third, it examines scholarship on judicial efficiency and comparative evaluation systems; and finally, it proposes an India-specific regulatory framework.

The contribution is deliberately narrower than the familiar argument that tribunals must be independent. The Act now creates a statutory evaluator. The unresolved issue is the legal architecture of evaluation itself. A metric may appear administrative but can become adjudicatory pressure when connected to reappointment, disciplinary scrutiny, resource allocation or public ranking. The paper therefore treats data governance as part of judicial independence rather than as a purely managerial subject.

The Constitutional Baseline: Tribunal Independence Is Functional, Not Formal

The Supreme Court's modern tribunal doctrine rests on a functional proposition: when adjudicatory work is transferred from constitutional courts to tribunals, the substitute institution must possess independence sufficient for the judicial character of the transferred function. In L. Chandra Kumar, tribunal decisions remained subject to High Court judicial review because the constitutional role of Articles 226 and 227 could not be excluded. In R. Gandhi and later Madras Bar Association cases, the Court extended

the concern beyond formal appellate review to the composition and institutional conditions of tribunals.

The November 2025 Madras Bar Association judgment is especially relevant. It reiterated that dependence on sponsoring departments for funding, infrastructure and personnel undermines independence and recorded the Court's long-standing concern that tribunals should resemble courts rather than bureaucratic boards.3 It then directed the Union to establish a National Tribunals Commission with independence from executive control, professional expertise, transparent processes and oversight mechanisms capable of reinforcing public confidence [4].

The constitutional value protected by these decisions is not merely freedom from explicit instructions about how to decide a case. Judicial independence also protects the decisional environment: tenure, promotion, disciplinary arrangements, workload design and institutional incentives must not reasonably create pressure to decide in a manner favourable to the evaluator. Performance measurement therefore falls within the same constitutional field whenever it can influence professional security or adjudicative behaviour.

1. Independence from Indirect Incentives

Tribunal independence is vulnerable not only to direct instructions but also to incentive structures that make future office depend on present adjudicatory behaviour. The 2026 framework is particularly sensitive because reappointment remains possible and the Government's own explanatory material states that previous service and performance may be considered in the selection process. A performance system that feeds unfiltered member-level data into reappointment therefore creates a risk structurally similar to the short-tenure problem identified in the Madras Bar Association line of cases: an adjudicator may formally remain free while reasonably perceiving that future tenure depends on satisfying an evaluator [9].

The constitutional inquiry should consequently examine three things together: the subject of measurement, the level at which the information is collected, and the consequence attached to it. Institutional data on vacancies or filing access pose little threat to decisional freedom. The same cannot be said of a member-specific 'performance score' that combines disposal, reversal and outcome data and is then supplied to a selection committee. The distinction is not between data and no data; it is between managerial information and information capable of operating as an indirect condition on judicial choice.

Why Performance Measurement Is Nevertheless Necessary

Independence cannot become a defence of institutional opacity. The Standing Committee on Personnel, Public Grievances, Law and Justice reported in August 2026 that several tribunals continued to experience vacancies and capacity constraints; for example, 27 of 126 sanctioned member posts in the Income Tax Appellate Tribunal were vacant as of May 2026 [5]. The very justification for specialist adjudication—speed, expertise and accessible remedies—is defeated when vacancies, listing delay and weak infrastructure reproduce ordinary-court backlog.

Empirical scholarship also shows why simplistic measurement is dangerous. Indian studies of court efficiency have used data-envelopment analysis and multi-factor frameworks to demonstrate that judicial performance depends on resources, staff, procedure, caseload complexity and exogenous factors rather than raw disposal numbers alone.6 Recent international research similarly cautions that clearance rates and disposition time, while useful, can provide an incomplete or misleading picture of efficiency if resource utilisation and case complexity are ignored [7].

The policy problem is therefore not whether to measure. It is what to measure, at what level of aggregation, with what contextual information, and with what legal consequences.

1. The Attribution Problem: Who Actually Causes Delay?

A credible performance regime must distinguish delay attributable to the tribunal from delay caused by the surrounding State apparatus. A bench cannot decide cases that have not been listed because registry posts are vacant; nor can it compensate indefinitely for delayed appointments, inadequate courtrooms, missing technical staff or non-functional digital infrastructure. If these structural constraints are folded into a single tribunal 'score', the State can statistically transfer responsibility for its own failures to the adjudicatory institution.

The NTC's annual reporting should therefore include an attribution statement for major indicators. Pending-case age, for example, should be disaggregated by periods of member vacancy, stayed proceedings, remands, statutory preconditions and other identifiable external causes. This is not an argument for excusing avoidable delay. It is a demand that accountability follow control. The institution capable of remedying a bottleneck should be the institution whose performance record reflects that bottleneck [10].

Comparative Lessons: Measuring Systems without Ranking Judges

The European Commission for the Efficiency of Justice (CEPEJ) offers a useful comparative model because its evaluation architecture measures judicial systems through combined information about budgets, professionals, access to justice, efficiency, quality and technology, rather than reducing judicial quality to a single productivity score [8]. Its database and country profiles are designed primarily for institutional diagnosis and comparative policy learning.

The comparative lesson is twofold. First, system-level indicators—resources, age of pending cases, staffing, digital access, legal aid, time standards—can reveal administrative failure without making an inference about how an individual adjudicator should decide. Second, even apparently objective indicators require contextualisation. A tribunal dealing with high-volume routine disputes cannot sensibly be compared, case for case, with a tribunal deciding technically complex environmental, competition or corporate matters.

1. What Comparative Efficiency Indicators Actually Measure

CEPEJ's commonly used clearance-rate and disposition-time indicators illustrate both the value and the limits of quantitative judicial management. Clearance rate compares resolved cases with incoming cases and can show whether backlog is expanding or contracting. Disposition time estimates how long the existing stock of pending cases would take to clear at the current rate of resolution. CEPEJ itself cautions that disposition time is a theoretical indicator

rather than the actual average duration of proceedings. Read together, the measures are useful for identifying flow problems; read as stand-alone quality scores, they are much less reliable [11].

Comparative practice also supports case-weighting and dashboard design that recognises heterogeneous workloads. A technically complex competition appeal, a multi-party environmental proceeding and a routine service dispute may each count as one 'case', yet they consume very different adjudicatory resources. The Indian Data Grid should therefore avoid league tables based on unweighted disposals. A more defensible approach would group comparable case types, publish median and percentile information rather than a single rank, and record the resource inputs available to the tribunal during the measurement period [12].

The Hidden Constitutional Risk in Section 4(b)

Section 4(b) does not specify whether the NTC may review only each tribunal as an institution or also benches and individual members. Nor does it identify permissible indicators. That ambiguity becomes more important because the 2026 framework allows reappointment and because prior service and performance may naturally become relevant to future selection decisions. Even without an express statutory command, an ecosystem can emerge in which individual members understand that low disposal numbers, high reversal rates or politically inconvenient outcomes may affect future office.

Three types of metric are particularly problematic. First, outcome metrics—such as the percentage of cases decided for or against the Union—would directly create a perception that substantive results are being monitored. Second, reversal-rate metrics can punish independent reasoning: a tribunal member may rationally choose doctrinal conformity over a good-faith interpretation if reversal carries career consequences. Third, crude productivity targets may create 'quantity pressure', encouraging premature hearings, shorter reasons or reluctance to grant procedural accommodations.

The danger is not hypothetical in organisational design. Metrics alter incentives. Once an indicator is tied to reputation, reappointment or institutional funding, participants optimise against it. In adjudication, that phenomenon can convert a management tool into indirect control over judicial method.

Proposed Independence-Compatible Performance Framework

The NTC should adopt a three-tier framework, preferably through publicly notified regulations or rules accompanied by a data-governance protocol. The categories should turn on the closeness of a metric to adjudicative substance and the severity of the consequence attached to it.

Category	Illustrative metrics	Permissible use
GREEN – institutional conditions	Vacancies; staff-to-member ratio; bench days lost; infrastructure; filing accessibility; digitisation; age profile of pending cases; time from vacancy to appointment	Collect, publish, benchmark, and use for resource allocation and administrative reform.
AMBER – process metrics	Median disposition time; adjournment frequency; reserved judgments pending beyond a defined period; clearance rate; listing delay	Use only with case-complexity controls, qualitative explanation and tribunal-level rather than automatic individual sanctions.
RED – decisional metrics	Government success rate; claimant/respondent win rate; ideological or subject-outcome patterns; reversal rate used as a career score; settlement pressure	Do not use for appointment, reappointment, discipline, remuneration, ranking or budgetary consequences.

Green metrics measure the capacity and administrative conditions within which adjudication occurs. They should form the core of the Data Grid. Amber metrics can identify delay, but only after controlling for complexity and procedural context. Red metrics intrude on the substance or perceived substance of adjudication and should be structurally firewalled from personnel decisions.

1. Consequence-Sensitive Use of Metrics

The legal risk of a metric depends as much on its use as on its content. Median disposition time is ordinarily a legitimate institutional indicator when used to identify staffing needs or ageing dockets. The same figure becomes more problematic if a member's reappointment turns automatically on beating a numerical target. Conversely, some member-level information—such as the period for which a reserved judgment remains undelivered—may be legitimately relevant where it reveals persistent non-performance after workload and health circumstances are considered.

For that reason, the Green-Amber-Red framework should be read as a matrix of both data type and consequence. Green data may support publication, benchmarking and resource allocation. Amber data may support diagnosis and, in exceptional cases, member-level review, but only through a reasoned contextual process. Red data may be collected in anonymised form for research or doctrinal study, but should be excluded from appointment, reappointment, discipline, remuneration and budget decisions. This consequence-sensitive approach is more precise than a blanket ban on judicial statistics.

Data Governance for the National Tribunals Data Grid

The Data Grid should not merely replicate a dashboard. Its legal design should incorporate purpose limitation. Data collected for transparency and system planning should not automatically migrate into disciplinary or selection processes. Access rights should differ between public institutional data, confidential case-management data and sensitive member-level analytics.

At least five safeguards are required: (1) a published data dictionary defining every metric; (2) prohibition on undisclosed composite scores; (3) case-complexity adjustment before cross-tribunal comparisons; (4) an opportunity for the concerned tribunal to append contextual explanations to annual performance findings; and (5) separation between the team maintaining the Data Grid and bodies making complaints or reappointment assessments. These safeguards transform data collection from managerial surveillance into accountable institutional diagnosis.

1. Auditability, Error Correction and Methodological Transparency

Because performance data can affect institutional reputation and future resource allocation, tribunals should be able to challenge the accuracy of the underlying dataset before annual findings are finalised. The NTC should maintain a documented correction process for duplicate matters, wrongly coded disposals, transferred cases and periods during which benches could not function. Methodological changes—such as a new case-weighting formula—should be versioned and accompanied by an explanation of their likely effect on time-series comparisons.

The public-facing Data Grid should also distinguish raw data from evaluative conclusions. Transparency is not served by publishing a composite score whose weighting formula is unknown. The more consequential the inference drawn from data, the stronger the obligation to disclose the definition, denominator, exclusions and limitations of the metric. In that sense, methodological transparency is part of natural justice at the institutional level: an adjudicatory body should not be publicly characterised as inefficient by a model it cannot understand or contest [11].

Reappointment, Complaints and the Firewall Principle

The most important doctrinal safeguard should be a 'firewall principle': data that reflects substantive decisional patterns must not be considered in reappointment or complaint proceedings. Legitimate reappointment material can include integrity findings, proven misconduct, chronic unexplained delay in delivering reserved decisions and persistent failure to discharge assigned work after contextual review. It should not include whether an adjudicator's decisions are frequently adverse to the government, commercially unpopular, or reversed on appeal.

Appellate reversal is a normal feature of adjudication. Treating reversal as evidence of poor performance collapses the distinction between appellate correction and disciplinary evaluation. It can also create vertical conformity inconsistent with adjudicatory independence. Any use of appellate data should therefore be aggregated for legal-training and systemic-learning purposes, not converted into a score for individual members.

Recommendations

- Rules under the Tribunals Reforms Act should expressly define “performance review” as institutional and administrative evaluation, not review of the merits of adjudicatory outcomes.
- The NTC should adopt and publish a Green–Amber–Red metric classification before the National Tribunals Data Grid becomes operational at scale.
- Individual-member dashboards, if maintained for workload management, should be confidential and barred from using party-success rates, ideological outcome measures or raw reversal rates.
- Every quantitative efficiency measure should be accompanied by a case-complexity and resource-context statement.
- Reappointment and complaint procedures should contain an express decisional-independence firewall preventing reliance on lawful adjudicatory outcomes.
- Annual NTC reports should separately report structural deficits attributable to the Union—vacancies, delayed appointments, infrastructure and staffing—so that tribunal members are not made statistically responsible for executive failures.

Conclusion

The National Tribunals Commission represents the most serious institutional attempt in decades to respond to the Supreme Court's demand for independent tribunal administration. Yet independence can be compromised not only by appointment control or financial dependence, but also by the metrics through which adjudicators are evaluated. Section 4(b) should therefore be implemented on the premise that efficiency data are constitutionally legitimate only when they measure the conditions and process of justice without rewarding particular adjudicatory outcomes.

A well-designed Data Grid can expose vacancies, delay, inadequate staffing and systemic bottlenecks and thereby strengthen the right to effective adjudication. A poorly designed performance regime can produce the opposite result: numerical incentives that silently reshape judicial behaviour. The constitutional task is to make tribunals accountable for administration without making adjudicators accountable for deciding cases in a preferred way. The proposed Green–Amber–Red framework offers a workable line between those two objectives.

References

1. S.P. Sampath Kumar v Union of India. (1987) 1 SCC 124.
2. L. Chandra Kumar v Union of India. (1997) 3 SCC 261.
3. Union of India v R. Gandhi, President, Madras Bar Association. (2010) 11 SCC 1.
4. Madras Bar Association v Union of India. 2025 INSC 1330.
5. Tribunals Reforms Act 2026. s 4.
6. Law Commission of India. Assessment of Statutory Framework of Tribunals in India. Report No 272. 2017.
7. Department-related Parliamentary Standing Committee on Personnel, Public Grievances, Law and Justice. Review of Functioning of Tribunal System in the Country. 2026.
8. Gupta M, Bolia NB. Efficiency Measurement of Indian High Courts Using DEA: A Policy Perspective. Journal of Policy Modeling,2020;42:1372.
9. Gupta M, Bolia NB. Factors Affecting Efficient Discharge of Judicial Functions: Insights from Indian Courts. Socio-Economic Planning Sciences,2024;91:101755.
10. Pereira MA. Clearance Rates and Disposition Times: Not the Whole Story of Judicial Efficiency. International Review of Law and Economics,2025;83:106283.
11. European Commission for the Efficiency of Justice (CEPEJ). Evaluation of Judicial Systems and CEPEJ-STAT.
12. Press Information Bureau. Reforming India’s Tribunal System. 12 Aug 2026.
13. European Commission for the Efficiency of Justice (CEPEJ). Guidelines and tools on judicial time management, court dashboards and case weighting.

14. Madras Bar Association v Union of India. 2025 INSC 1330. Supreme Court of India; 2025 Nov 19. Available from: https://indiankanoon.org/doc/17580897/
15. Tribunals Reforms Act 2026. s 4; Press Information Bureau. Reforming India's Tribunal System. 12 Aug 2026. Available from: https://www.pib.gov.in/PressReleasePage.aspx?PRID=2298157
16. Madras Bar Association v Union of India. 2025 INSC 1330. Discussion of Union of India v R Gandhi and dependence on sponsoring departments. Available from: https://indiankanoon.org/doc/17580897/
17. Madras Bar Association v Union of India. 2025 INSC 1330. Paras 152–153.
18. Department-related Parliamentary Standing Committee on Personnel, Public Grievances, Law and Justice. Review of Functioning of Tribunal System in the Country. 7 Aug 2026. Available from: https://prsindia.org/policy/report-summaries/review-of-functioning-of-tribunal-system-in-the-country
19. European Commission for the Efficiency of Justice (CEPEJ). Evaluation of Judicial Systems. 2026 cycle. Available from: https://www.coe.int/en/web/cepej/cepej-work/evaluation-of-judicial-systems-2026
20. Press Information Bureau. Reforming India's Tribunal System. 12 Aug 2026. Available from: https://www.pib.gov.in/PressReleasePage.aspx?PRID=2298157
21. European Commission for the Efficiency of Justice (CEPEJ). Guidelines on judicial time management and performance indicators. Available from: https://rm.coe.int/cepej-guidelines/16809f007a
22. European Commission for the Efficiency of Justice (CEPEJ). CEPEJ-STAT. Available from: https://www.coe.int/en/web/cepej/cepej-stat
23. European Commission for the Efficiency of Justice (CEPEJ). Tools on judicial time management: Handbook on Court Dashboards and Study on Case Weighting in Judicial Systems. Available from: https://www.coe.int/en/web/cepej/saturn-tools
24. Law Commission of India. Assessment of Statutory Framework of Tribunals in India. Report No 272. 2017.