



DESEMPENHO LABORATORIAL FORENSE: proposta de uma estrutura de avaliação multinível



Adelino Pinheiro Silva

<http://lattes.cnpq.br/8373538496107754> – <https://orcid.org/0000-0002-2796-4841>
adelino.pinheiro@policiacivil.mg.gov.br

Polícia Civil de Minas Gerais, Belo Horizonte, MG, Brasil

RESUMO

O Brasil enfrenta uma crise persistente de segurança pública, com mais de 47 mil mortes violentas intencionais anuais e taxa de elucidação de homicídios entre 20% e 40%. Nessa crise, os laboratórios forenses operam com recursos limitados, *backlogs* (pendências) crescentes e pressões por produtividade. A aplicação incondicional do paradigma gerencialista no contexto forense tem como limitação as métricas operacionais, como *turnaround time* (TAT - tempo entre a requisição e a expedição do laudo) e produtividade, e trata como isolados sintomas de dinâmicas criminológicas multifatoriais e exógenas aos laboratórios. Diante desse cenário, o presente trabalho tem por objetivo propor uma estrutura de avaliação multinível que concilie a eficiência com a consciência de que a demanda pericial deriva de fatores estruturais que escapam ao controle laboratorial. Mediante revisão da literatura, sistematizaram-se indicadores em quatro níveis de análise: macro (contexto criminológico e sistêmico), meso (interface institucional), micro (desempenho laboratorial) e integração. A estrutura especifica 32 indicadores distribuídos nos três primeiros níveis e propõe o Índice de Desempenho Laboratorial Forense (IDLF), métrica sintética que agrega as dimensões em escala de 0 a 1, com capacidade de decomposição para diagnósticos granulares. Visando à melhoria dos serviços públicos, sugere-se a interoperabilidade entre sistemas de informação, o investimento em tecnologias de *big data* e inteligência artificial, bem como programas de proficiência.

Palavras-chave: Laboratório forense; Avaliação de desempenho; Indicadores multinível; Gerencialismo; Sistema de justiça criminal.

FORENSIC LABORATORY PERFORMANCE: proposal for a multilevel assessment framework

ABSTRACT

Brazil faces a persistent public security crisis, with over 47,000 intentional violent deaths annually and homicide clearance rates between 20% and 40%. In this context, forensic laboratories operate with limited resources, increasing backlogs, and productivity pressures. The unconditional application of the managerialist paradigm in the forensic context has a key limitation: operational metrics such as turnaround time and productivity treat as isolated issues what are actually symptoms of multifactorial criminological dynamics exogenous to laboratories. Given this scenario, this paper aims to propose a multilevel assessment framework that reconciles efficiency with the awareness that forensic demand derives from structural factors beyond laboratory control. Through a literature review, indicators were systematized across four levels of analysis: macro (criminological and systemic context), meso (institutional interface), micro (laboratory performance), and integration. The framework specifies 32 indicators distributed across the first three levels and proposes the Forensic Laboratory Performance Index (FLPI), a synthetic metric that aggregates dimensions on a scale from 0 to 1, with decomposition capability for granular diagnostics. Aiming at public service improvement, interoperability among information systems, investment in big data and artificial intelligence technologies, as well as proficiency programs are recommended. It is acknowledged that improving forensic services entails shared responsibility among multiple institutional actors.

Keywords: Forensic laboratory; Performance assessment; Multilevel indicators; Managerialism; Criminal justice system.

DOI: <https://doi.org/10.70365/2764-0779.2026.184>

Recebido em: 09/02/2026.

Aceito em: 26/03/2025.

1 INTRODUÇÃO

O Brasil enfrenta uma crise persistente de segurança pública caracterizada por elevadas taxas de criminalidade violenta, baixa efetividade do Sistema de Justiça Criminal (SJC) e crescente demanda por serviços relacionados à segurança pública. Os dados do Fórum Brasileiro de Segurança Pública (FBSP, 2025) registram mais de 44 mil mortes violentas intencionais (MVI) anuais – uma taxa de 20,8 por 100 mil habitantes, que reduziu 5,4% em relação ao ano anterior – que ainda é uma das maiores taxas do mundo. Segundo a análise do Banco Mundial (2018), o SJC brasileiro tem combinado tendências de encarceramento e incapacidade de prevenir crimes e de punir infratores de maneira coerente.

Nesse contexto, os serviços forenses ocupam posição estratégica na integração da polícia judiciária, sendo que a prova técnica pode agregar robustez à investigação policial e contribuir na fundamentação de denúncias ministeriais e subsidiar decisões judiciais. Contudo, os laboratórios forenses brasileiros operam, frequentemente, com recursos limitados, *backlogs* crescentes e pressões por maior produtividade (Senasp, 2013). Na mesma tendência, a taxa de elucidação de homicídios no Brasil situa-se entre 20% e 40%, dependendo da jurisdição, inferior aos 60-90% observados em países desenvolvidos (Costa, 2015; Ribeiro & Diniz, 2020).

A literatura internacional sobre gestão de laboratórios forenses consolidou-se nas últimas duas décadas em torno de estruturas de avaliação quantitativa bem estabelecidas. O projeto FORESIGHT¹ tem-se consolidado como padrão internacional para avaliação de desempenho, cobrindo mais de 200 laboratórios em 27 países com dados longitudinais de 12 anos (Houck & Speaker, 2022). Os indicadores padronizados incluem: (i) *turnaround time* (TAT), que é o tempo entre a requisição e a expedição do laudo; (ii) custo por caso; (iii) casos por analista em *full-time equivalent*² (FTE); e (iv) taxa de *backlog*, que é a proporção de pendências em relação a um período de tempo fixo.

¹ O Project FORESIGHT é uma iniciativa de autoavaliação e *benchmarking* baseada em métricas de negócios, voltada para laboratórios forenses em âmbito global. Criado em 2008, o projeto consolidou-se como referência na avaliação de serviços forenses. A iniciativa utiliza definições padronizadas para diversas áreas forenses, permitindo comparação pareada entre os laboratórios participantes e outras instituições acreditadas (ISO/IEC 17025). O projeto coleta e analisa dados anuais sobre volume de casos, alocação de pessoal, tempo de resposta e despesas financeiras. O objetivo do FORESIGHT é auxiliar os gestores laboratoriais na avaliação da eficiência operacional, na otimização de recursos e na demonstração do valor dos serviços forenses para o sistema de justiça criminal.

² O *full-time equivalent* (FTE), ou equivalente em tempo integral, é uma métrica padronizada para quantificar a força de trabalho. Essa unidade converte as horas de trabalhadores em tempo integral, meio período e horas extras em uma medida uniforme. No contexto forense, o FTE permite estimar índices de produtividade, dimensionar a capacidade laboral e viabiliza comparações entre diferentes instituições.

As avaliações comparativas do projeto FORESIGHT revelam, entre 2011-2017, aumento de 60% no TAT e crescimento de 250% no *backlog* de 30 dias (Houck & Speaker, 2022). Na mesma tendência, Speaker (2021) estimou a necessidade de pessoal frente à demanda de laboratórios norte-americanos utilizando regressões não-lineares. Os resultados apontaram para um déficit superior a 900 profissionais, com os laboratórios forenses operando, em média, 19,7% abaixo da capacidade calculada como ótima.

No âmbito metodológico, Omari e colaboradores (2018) propuseram decomposição do TAT em seis componentes mensuráveis, permitindo a identificação de gargalos específicos mediante análise de Pareto. Alfaro e colaboradores (2020) documentaram a aplicação da metodologia *lean six sigma* na unidade de Balística da Costa Rica, alcançando uma redução de 75% no TAT e 97% nos casos pendentes. Essas ações metodológicas indicam que é possível mensurar processos em laboratórios forenses e que pode ser viável otimizá-los.

Frente a esse indicativo, a aplicação do paradigma gerencialista em laboratórios forenses deriva da reforma administrativa iniciada nas décadas de 1980-1990 (Hood, 1991). Dentre os componentes básicos desse paradigma, incluem a gestão profissional, os padrões explícitos de desempenho, a ênfase no controle de resultados, e a disciplina no uso de recursos (Hood, 1995; Pollitt, 1990). No Brasil, a reforma administrativa de 1995 incorporou estes princípios ao setor público, inclusive na segurança pública (Bresser-Pereira, 1996). Apesar dos avanços em métricas operacionais, a literatura apresenta lacunas quanto à articulação entre desempenho laboratorial e efetividade sistêmica.

Entretanto, no contexto forense, o paradigma gerencialista não pode desconsiderar três aspectos: (i) a demanda por serviços periciais deriva de fatores criminológicos exógenos; (ii) a eficiência operacional não garante efetividade sistêmica, pois laudos podem ser subutilizados nas etapas subsequentes do fluxo criminal; (iii) o plano nacional de segurança pública não integra diferentes níveis de análise. Os modelos disponíveis tratam cada nível isoladamente, comprometendo diagnósticos contextualizados.

Frente ao contexto exposto, o presente trabalho tem como objetivo propor uma estrutura de avaliação multinível de indicadores de desempenho para laboratórios forenses que concilie a necessidade de gestão eficiente com o reconhecimento de que a demanda pericial deriva de dinâmicas criminológicas multifatoriais e exógenas. Mais especificamente, o trabalho busca revisar os trabalhos prévios visando elencar metodologias bem estabelecidas, sistematizar indicadores em quatro níveis de análise (macro, meso, micro, integração), propondo suas fontes, periodicidade e

interpretação, e formular recomendações que considerem tanto a eficiência operacional quanto a efetividade sistêmica.

A contribuição teórica reside na conciliação de perspectivas do gerencialismo e da criminologia através da proposta que reconhece a legitimidade de ambas as perspectivas. A contribuição prática oferece instrumentos de diagnóstico e gestão aplicáveis a laboratórios forenses brasileiros, com potencial de generalização para outros contextos. A estrutura de avaliação fundamenta-se em dois princípios metodológicos complementares: o primeiro, frequentemente associado à revolução científica – embora falsamente atribuído a Galileu (Kleinert, 2009) –, sustenta a necessidade de “medir o que é mensurável e tornar mensurável o que não é” ; e o segundo, o adágio gerencial (“o que é medido é gerenciado”), deriva do alerta de Ridgway (1956) sobre disfunções da medição, reforçando a necessidade de interpretação cautelosa dos indicadores. A tentativa de estender esses princípios desde o nível macro, no âmbito da segurança pública, até o nível micro, na gestão de um laboratório forense, constitui a base metodológica deste trabalho.

O presente trabalho, organizado em cinco seções, apresenta a fundamentação teórica na próxima seção, abordando o paradigma gerencialista, o SJC brasileiro, e a distinção entre eficiência e efetividade. A seção 3 detalha a estrutura de avaliação multinível proposta, especificando indicadores para cada nível. A seção 4, por sua vez, discute as implicações teóricas e práticas da proposta, e a quinta seção apresenta as considerações finais, as recomendações e as propostas de continuidade.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 Nova Gestão Pública e Gerencialismo

A Nova Gestão Pública (NGP) emergiu como proposta para agregar práticas do setor privado à administração pública. Hood (1991) identificou sete componentes doutrinários centrais: gestão profissional ativa, padrões explícitos de desempenho, ênfase no controle de resultados, desagregação de unidades, competição no setor público, estilos de gestão privada e disciplina no uso de recursos. O autor distinguiu tensões entre *sigma-values* (eficiência), *theta-values* (equidade) e *lambda-values* (resiliência), observando que a NGP enfatiza os primeiros em detrimento dos demais. Essa tensão é particularmente relevante para serviços de Justiça, nos quais equidade e legitimidade processual podem conflitar com pressões por eficiência.

Pollitt (1990) ofereceu análise crítica do gerencialismo, argumentando que a transposição acrítica de técnicas privadas ignora especificidades do setor público: ausência de mecanismo de preços, multiplicidade de objetivos conflitantes e natureza não-voluntária de muitos serviços. No caso de laboratórios forenses, o "cliente" não escolhe o serviço, não paga diretamente, e o "produto" (laudo pericial) integra processo decisório coletivo envolvendo múltiplas instituições.

Na segurança pública, o gerencialismo manifestou-se em iniciativas como o Compstat³, um modelo de gestão que enfatiza inteligência oportuna, táticas efetivas, implementação rápida de recursos e acompanhamento rigoroso (Weisburd *et al.*, 2003). Entretanto, a literatura documenta efeitos adversos. Eterno e Silverman (2012) apontaram manipulação de estatísticas em departamentos sob pressão por indicadores. Manning (2001) argumentou que o modelo privilegia métricas de processo sobre resultados substantivos. Para laboratórios forenses, riscos análogos incluem priorização de casos simples para melhorar TAT, pressão por identificações positivas comprometendo critérios técnicos e negligência de atividades não-mensuráveis, como pesquisa metodológica.

No Brasil, o Plano Diretor da Reforma do Aparelho do Estado (PDRAE) propôs transição para administração gerencial com foco em resultados (Bresser-Pereira, 1996). Os serviços forenses situam-se nas "atividades exclusivas" do Estado, para as quais o Plano previa contratos de gestão baseados em indicadores de desempenho. Embora a implementação tenha sido parcial, o paradigma gerencialista permeia a administração pública brasileira contemporânea.

2.2 Nova Gestão Pública e Gerencialismo

O SJC brasileiro compreende as instituições responsáveis pelo processamento de eventos criminais: Polícias Militar e Civil, Ministério Público, Poder Judiciário, Defensoria Pública e sistema penitenciário. A fragmentação institucional é característica distintiva: a Constituição de 1988 estabeleceu ciclo incompleto de polícia, criando interfaces que impactam a celeridade do sistema (Sapori, 2007). A perícia criminal insere-se no subsistema de investigação, podendo vincular-se à Polícia Civil, à Secretaria de Segurança Pública ou constituir órgão autônomo. Essa heterogeneidade dificulta análises comparativas e padronização de procedimentos (Senasp, 2013).

³ O Compstat (Computer Statistics) é um modelo de gestão policial introduzido em 1994 no Departamento de Polícia de Nova York, focado no uso de dados e na responsabilização para o controle da criminalidade.

As pesquisas sobre fluxo do SJC reconstituem o processamento de eventos criminais por meio das instâncias. Vargas (2004, 2014) sistematizou as dificuldades metodológicas: fragmentação de bases de dados, inexistência de identificadores únicos, descontinuidade de registros e heterogeneidade classificatória. Estudos empíricos revelam baixa efetividade. Ribeiro e Diniz (2020) mostraram que apenas uma fração dos homicídios registrados em Belo Horizonte resulta em condenação. Cano e Duarte (2010) estimaram que mais de 90% dos homicídios no Rio de Janeiro não resultam em responsabilização criminal. Costa (2015) encontrou taxa de condenação de 32,4% no Distrito Federal, valor superior a outras capitais, mas inferior a países desenvolvidos.

O Banco Mundial (2018) caracterizou o sistema como "esticado ao limite", combinando encarceramento elevado com baixa capacidade de prevenção e dissuasão. Cook (1979) pontuou que a taxa de elucidação constitui medida imperfeita de efetividade policial, pois não captura uma qualidade das investigações nem tradução em condenações. Borg e Parker (2001) mostraram que a estrutura social urbana influencia taxas de elucidação, sugerindo que variações entre jurisdições refletem fatores além da competência investigativa. No Brasil, a taxa de elucidação de homicídios situa-se entre 20% e 40% (Ipea, 2016) enquanto, em Minas Gerais, a taxa de elucidação de homicídios, em 2023, foi de 45,2% (Minas Gerais, 2024, p. 61). No conjunto, os autores analisados apontam para causas como a sobrecarga investigativa, morosidade judicial, desarticulação institucional e escassez de recursos

2.3 Eficiência versus Efetividade

A distinção entre eficiência e efetividade permanece central na administração pública (Rutgers & van der Meer, 2010; Meier *et al.*, 2023). Eficiência representa a razão entre insumos e produtos; efetividade relaciona produtos a resultados sociais desejados (Mandl *et al.*, 2008). No caso laboratorial, eficiência refere-se a métricas como TAT, casos por FTE e taxa de conclusão, mensuráveis com dados administrativos. As avaliações do projeto FORESIGHT operacionalizam primordialmente métricas de eficiência.

Efetividade refere-se ao impacto na resolução do problema que justifica a existência do serviço, por exemplo, a contribuição para a elucidação de crimes, sentenças justas e redução da criminalidade. Mandl e colaboradores (2008) enfatizam que eficiência operacional não garante efetividade sistêmica. Um laboratório pode apresentar excelentes indicadores de TAT enquanto laudos pouco conclusivos permanecem subutilizados nas etapas subsequentes do fluxo criminal. Inversamente, um laboratório com indicadores

modestos pode ter alta efetividade se seus laudos forem decisivos para elucidação de crimes graves.

Essa distinção fundamenta a necessidade de uma estrutura de avaliação que articule métricas operacionais (eficiência) com indicadores de resultado sistêmico (efetividade). Essa articulação busca conciliar o paradigma gerencialista – sem ignorar efetividade em favor de eficiência –, com o impasse analítico decorrente da complexidade sistêmica – sem abandonar métricas operacionais por dificuldade de medir impacto final.

2.4 Avaliação em Laboratórios Forenses: Projeto FORESIGHT

A necessidade de avaliação padronizada foi identificada nos EUA, na década de 1960. Parker (1963) constatou que a ajuda científica na investigação criminal era "ferramenta de utilidade negligenciável", sem uso de evidência científica em mais de 98% das violações criminais conhecidas. Benson, Stacy e Worley (1970) documentaram uma heterogeneidade nas práticas que impedia comparações: a depender do laboratório, seis amostras de caligrafia podiam ser registradas como seis exames ou como um único. Os autores concluíram que o envolvimento do laboratório forense no corpo total de crimes era "tão minúsculo que impede um julgamento quanto ao impacto da criminalística no sistema de justiça criminal". Essa lacuna metodológica persistiria por quatro décadas até a criação do Projeto FORESIGHT. A Tabela 1 apresenta a evolução histórica dos custos desde os primeiros levantamentos sistemáticos.

O Projeto FORESIGHT iniciou uma coleta sistemática em 2008-2009 e tem-se consolidado com coleta de dados longitudinais de 200 laboratórios em 27 países. A metodologia estabeleceu definições padronizadas para 21 áreas forenses em dois níveis (Houck & Speaker, 2020; 2024). O Nível I inclui carga de trabalho, pessoal e despesas. O Nível II expande com informações sobre itens examinados, TAT e *backlog* detalhado (Speaker, 2021). As definições operacionais padronizam os termos como: caso (requisição que resulta em exames), FTE (equivalente a 2.080 horas/ano), *backlog* (casos não completados ao final do período) e TAT (tempo entre a requisição e a expedição do laudo).

Tabela 1 - Evolução dos custos de laboratórios forenses (1963-2019). Valores de 1963 e 1968 ajustados para dólares de 2019.

Período	Custo por caso (US\$)	Custo por 100 mil habitantes (mín.; mediana; máx; em US\$)
1963 (Parker) ajustado	184 – 426 – 919	17.535; 61.874; 200.400
1968 (OLEA) ajustado	186 – 465 – 930	17.595; 62.086; 201.089
2019 (FORESIGHT)	916 – 1.411 – 2.450	245.114; 360.431; 671.114
Aumento 1968→2019	+392% – +203% – +163%	+1.293%; +481%; +234%

Fonte: Houck & Speaker (2020), compilando dados de Parker (1963) e OLEA (1968).

Os indicadores principais incluem custo por caso (orçamento dividido por casos completados), produtividade (casos por FTE em função-fim), taxa de *backlog* (proporção de casos pendentes há mais de 30 dias) e custo por 100 mil habitantes. A Tabela 2 apresenta os valores consolidados e tendências preocupantes na década 2011-2019.

Em paralelo ao projeto FORESIGHT, Speaker (2021) desenvolveu uma ferramenta econométrica para estimar necessidades de pessoal, considerando tipo de jurisdição, população e taxas de criminalidade. Como já colocado, os achados indicam que laboratórios estadunidenses operam com 19,7% menos pessoal que o necessário para eficiência ótima. A contribuição de Omari e colaboradores (2018) separou o TAT em seis etapas: processos administrativos iniciais (TAT1, 2-5 dias), análise técnica (TAT2, 80-95 dias), redação (TAT3, 1-10 dias), revisão (TAT4, 3-13 dias), aprovação técnica (TAT5, 9-26 dias) e aguardo de retirada pelo requisitante (TAT6, 19-27 dias). A análise de Pareto identificou que TAT2 contribui com 65-70% do tempo total, seguido por TAT6 com 15-20%. Essa decomposição permitiu investir em capacidade analítica e técnica apontando-o como gargalo para redução do TAT.

Tabela 2 - Os principais indicadores consolidados pelo FORESIGHT revelam tendências na década 2011-2019.

Indicador	Mediana 2019	Tendência 2011-2019
Custo por caso	US\$ 1.411	+15%
Casos por FTE/ano	95	-8%
TAT mediano	46 dias	+60%
Backlog 30 dias	15%	+250%
Custo por 100k habitantes	US\$ 360.431	—

Fonte: Adaptado de Houck & Speaker (2020); Speaker (2021); dados agregados do *Project FORESIGHT Annual Reports*.

Um achado metodológico refere-se à elasticidade de demanda negativa de -1,29, ou seja, para cada caso em que o *backlog* é reduzido, aproximadamente 1,29 novos casos são submetidos (Speaker, 2021). Uma das hipóteses desse efeito seria que, quando o TAT é elevado, requisitantes deixam de submeter exames aguardando resultados pendentes. Por outro lado, quando o laboratório melhora sua capacidade, essa demanda reprimida materializa-se. A implicação prática é que reduzir TAT sem aumentar proporcionalmente a capacidade pode paradoxalmente aumentar o *backlog* no curto prazo.

Apesar dos achados, o FORESIGHT tem limitações. Houck & Speaker (2020) apontam a possibilidade de viés de participação, em que laboratórios participantes seriam mais engajados em gestão de qualidade. A heterogeneidade jurisdicional limita comparabilidade direta. Estruturalmente, o projeto mede a eficiência operacional, mas não o impacto na justiça criminal. Conforme advertem os autores, a avaliação possibilita o diagnóstico e a comparação, mas não garante a melhoria sem implementação ativa, uma vez que "o setor privado tem lucro como indicador único de efetividade; o setor público carece de indicador equivalente" (Houck & Speaker, 2020).

A advertência de Goodhart (1984) de que "quando uma medida torna-se meta, deixa de ser boa medida" expõe uma contradição prática do gerencialismo analisado por Pollitt (1990). Ao impor metas de produtividade (como o TAT) próprias do setor privado, incentiva-se, no paradoxal contexto forense, que se produzam mais laudos em menos tempo, mas não necessariamente melhores ou mais confiáveis, esvaziando assim a própria noção de eficiência no serviço público. Isso posto, a aplicação do FORESIGHT ao contexto brasileiro requer adaptações metodológicas específicas, incluindo conversão monetária, estrutura jurisdicional (predominância de laboratórios estaduais), regime de trabalho (1.760 a 2.080 horas/ano) e marco regulatório (ISO 17025/ABNT versus ASCLD). A escassez de dados longitudinais padronizados no Brasil permanece como obstáculo crítico (Senasp, 2013), embora sistemas como o Geia (Mato Grosso) e o PCnet (Minas Gerais) ofereçam infraestrutura potencial para coleta padronizada.

3 ESTRUTURA DE AVALIAÇÃO MULTINÍVEL PROPOSTA

A presente seção apresenta os indicadores nos níveis de análise macro, meso e micro, incluindo fontes, interpretação e limitações. A principal crítica metodológica aos indicadores sistêmicos é a dificuldade de atribuição causal: a contribuição dos laboratórios forenses para taxas de elucidação, denúncia

ou condenação é mediada por múltiplos fatores, tornando inadequada a utilização isolada dessas métricas para avaliação do desempenho pericial. A estrutura proposta complementa indicadores contextuais com métricas de impacto específicas no nível micro e indicadores de integração no nível quatro.

3.1 Nível Macro: Contexto Criminológico e Sistêmico

O primeiro nível situa o laboratório forense em contexto mais amplo, assumindo que a demanda por serviços periciais deriva de dinâmicas criminológicas e socioeconômicas que fogem do controle do próprio laboratório. Essa perspectiva fundamenta-se na literatura criminológica, que demonstra correlação entre indicadores estruturais de desigualdade e taxas de criminalidade (Cerqueira & Lobão, 2004; Beato, 2012).

A inclusão de indicadores macro complementa abordagens gerencialistas. Se um laboratório apresenta *backlog* de 30 dias crescente (proporção de casos pendentes há mais de 30 dias), a análise pontual tenderia a atribuir o fenômeno a ineficiências internas. Uma estrutura multinível agrega ao diagnóstico a verificação de condições contextuais, como variações nas taxas de criminalidade, mudanças nos padrões de registro policial ou alterações na política de requisição de perícias, fatores que podem explicar total ou parcialmente o aumento de demanda. Dessa forma, propõem-se indicadores macro, os estruturais, de criminalidade e os sistêmicos.

Como indicadores estruturais, sugere-se o Coeficiente de Gini, que mensura desigualdade de renda em escala de zero (igualdade) a um (desigualdade máxima). A literatura documenta correlação entre desigualdade e criminalidade violenta (Fajnzylber, Lederman & Loayza, 2002), embora a relação causal permaneça em debate. A limitação principal é seu caráter agregado, que mascara heterogeneidades locais relevantes para jurisdições específicas. A Taxa de Desemprego Juvenil (15-24 anos) fundamenta-se na teoria das atividades rotineiras (Cohen & Felson, 1979) e em estudos brasileiros (Cerqueira, 2014) que identificam jovens desempregados como grupo mais exposto à vitimização e ao envolvimento em atividades ilícitas. O IDH (Índice de Desenvolvimento Humano) sintetiza saúde, educação e renda mediante média geométrica (Sen, 2000). O Instituto de Pesquisa Econômica Aplicada (IPEA) disponibiliza a versão municipal (IDHM), permitindo análises territoriais desagregadas. Sua limitação é a periodicidade decenal. O Índice de Violência Social (IVS) representa refinamento conceitual ao incorporar 16 indicadores em três dimensões: infraestrutura urbana, capital

humano e renda/trabalho (Costa & Marguti, 2015). Varia de 0 a 1, com valores elevados, indicando maior vulnerabilidade, e compartilha a limitação de periodicidade decenal.

Para indicadores de criminalidade, pode-se utilizar a Taxa de Mortes Violentas Intencionais (MVI), que constitui o indicador mais abrangente de letalidade violenta no Brasil, agregando: crimes violentos letais intencionais (CVLI, que inclui homicídios dolosos, latrocínios e lesões corporais seguidas de morte), mortes decorrentes de intervenção policial (MDIP) e mortes violentas com causa indeterminada (MVCI). O padrão internacional situa-se abaixo de 10 por 100 mil habitantes; o Brasil registrou 22,8 em 2022, variando entre 8,8 no Sul e 38,2 no Nordeste (UNODC, 2019; FBSP, 2023). A limitação desses indicadores encontra-se no fato de que mortes com causa indeterminada podem indicar homicídios ocultos que demandariam reclassificação posterior (Cerqueira, 2013). As limitações derivam da subnotificação, estimada entre 70% e 80% para crimes patrimoniais (Cardia *et al.*, 2003), e das diferenças classificatórias interjurisdicionais (Lima *et al.*, 2014).

Por fim, os indicadores sistêmicos sugeridos são a Taxa de Elucidação, a qual indica a proporção de inquéritos relatados com autoria identificada. Países desenvolvidos operam entre 60% e 90% para homicídios; no Brasil, estimativas indicam 20% a 40% (Costa, 2015; Ribeiro & Diniz, 2020). A baixa taxa pode relacionar-se tanto à contribuição pericial para identificação de autoria quanto à precarização do serviço que prejudica a polícia judiciária. Já a Taxa de Denúncia indica a proporção de inquéritos que evoluem para persecução penal pelo Ministério Público. A Taxa de Condenação indica a proporção de denúncias que resultam em condenação judicial. Juntas, permitem análise das etapas do SJC, identificando onde ocorrem as maiores perdas. O Tempo Médio de Processamento complementaria a análise. Dados do CNJ (2024, 2025) indicam, em média, 7 anos e 6 meses até arquivamento de ações penais de crimes dolosos contra a vida. A Tabela 3 resume os indicadores de nível macro, distinguindo as categorias, a fonte, a periodicidade e as principais limitações.

Tabela 3 – Resumo dos indicadores do nível macro de contexto criminológico e sistêmico.

Categoria	Indicador	Fonte	Periodicidade de	Limitações
Estruturais	Coeficiente de Gini	IBGE/PNAD	Anual	Caráter agregado
	Taxa desemprego juvenil (15-24)	PNAD Contínua	Trimestral	Não capta informalidade
	IDH Municipal	Ipea/Atlas	Decenal	Periodicidade limitada
	IVS	Ipea	Decenal	Periodicidade limitada
Criminalidade	Taxa MVI/100 mil hab.	FBSP	Anual	Subnotificação variável
	Taxa crimes patrimoniais	SSP/UF	Anual	70-80% subnotificação
Sistêmicos	Taxa de elucidação	Polícias Cíveis	Variável	Dados fragmentados
	Taxa de denúncia	MP	Variável	Integração precária
	Taxa de condenação	TJ	Variável	Difícil rastreamento
	Tempo médio processamento	CNJ	Anual	7,5 anos (média)

Fonte: Elaborado pelo autor com base em IBGE, Ipea, FBSP (2023), UNODC (2019), CNJ (2024; 2025), e literatura citada. **Notas:** IBGE: Instituto Brasileiro de Geografia e Estatística, PNAD: Pesquisa Nacional por Amostragem de Domicílios, Ipea: Instituto de Pesquisa Econômica Aplicada, IVS: Índice de Vulnerabilidade Social, MVI: Mortes Violentas Intencionais, FBSP: Fórum Brasileiro de Segurança Pública, SSP/UF: Secretaria de Segurança Pública das Unidades Federativas, MP: Ministério Público, TJ: Tribunal de Justiça, CNJ: Conselho Nacional de Justiça.

3.2 Nível Meso: Interface Institucional

O segundo nível busca indicadores da interface entre o laboratório forense e as demais instituições do SJC. A literatura sobre fluxo do sistema brasileiro indica gargalos e descontinuidades na passagem de casos entre polícia, perícia, Ministério Público e Judiciário (Vargas, 2014; Ribeiro & Silva, 2010). Nessas circunstâncias, propõem-se indicadores nas categorias de integração institucional, capacidade do SJC e de qualidade processual.

Como indicadores de integração institucional, pode-se citar o Tempo entre Ocorrência e Requisição, que pode comprometer, dependendo do vestígio, a identificação de indivíduos e a robustez probatória. A natureza do vestígio determina sua tolerância temporal. A limitação é que o indicador refere-se ao comportamento de agentes externos ao laboratório. A Taxa de Requisições Completas avalia a proporção que chega ao laboratório, atendendo todos os requisitos para exame. Requisições incompletas geram retrabalho administrativo, retardam análise técnica e podem resultar em devoluções que inflam estatísticas de produtividade. Baixa completude pode indicar necessidade de capacitação de requisitantes ou revisão de protocolos. A Taxa de Retornos para Complementação indica a frequência de solicitações adicionais após recebimento inicial. Diferentemente da métrica anterior, ela indica adequação da comunicação entre requisitante e

laboratório. Taxas elevadas podem indicar quesitação inadequada, materiais insuficientes ou expectativas desalinhadas. A Taxa de Contestações em Juízo pode indicar falhas metodológicas ou estratégias defensivas protelatórias. O indicador precisaria distinguir contestações procedentes (problemas de qualidade) de improcedentes (robustez técnica). A limitação é que a disponibilidade depende de acompanhamento processual que extrapola atribuições laboratoriais.

A capacidade do SJC pode ser inferida pela Razão Inquéritos por Delegado, a qual indica carga de trabalho policial. Dados da Senasp (2013) indicavam média nacional de 350, com variação de 150 a 800 entre estados. A heterogeneidade dos inquéritos pode distorcer o índice. A Razão Processos por Promotor e Processos por Juiz aplica lógica análoga às etapas ministerial e judicial. O Índice de Congestionamento do Judiciário, calculado pelo CNJ como proporção de casos pendentes em relação ao total, complementa a análise com perspectiva de estoque processual.

Para a qualidade processual pode-se medir a Taxa de Perícias em Relação aos Crimes, que indica a proporção de ocorrências que demandam análise técnica. Taxas baixas podem indicar subutilização do recurso pericial; taxas elevadas podem indicar requisições indiscriminadas que sobrecarregam o laboratório. O Tempo Fato-Expedição do Laudo indica tempestividade da requisição (responsabilidade externa) e celeridade laboratorial (responsabilidade interna). A decomposição em componentes externos e internos permite atribuição precisa de responsabilidades. A Tabela 4 sintetiza os indicadores de nível meso distinguindo as categorias, com uma breve definição e interpretação.

Tabela 4 – Resumo dos indicadores do nível meso de interface institucional.

Categoria	Indicador	Definição	Interpretação
Integração	Tempo ocorrência→requisição	Dias entre fato e chegada ao laboratório	Responsabilidade externa
	Taxa requisições completas	% com documentação adequada	Qualidade da interface
	Taxa retornos complementação	% que demanda informação adicional	Comunicação laboratório e requisitante
	Taxa contestações em juízo	% laudos contestados	Distinguir procedentes/improcedentes
Capacidade SJC	Inquéritos/delegado	Razão carga de trabalho	Média nacional: 350
	Processos/promotor	Razão MP	Variável por UF
	Índice congestionamento	CNJ: pendentes/total	Meta Judiciário
Qualidade	Taxa perícias/crimes	% ocorrências periciadas	Utilização recurso pericial
	Tempo fato→expedição laudo	Dias totais	Decompor em externo/interno

Fonte: Elaborado pelo autor com base em Senasp (2013), CNJ, Vargas (2014) e Ribeiro & Silva (2010). **Notas:** SJC: Sistema de Justiça Criminal, UF: Unidades Federativas, MP: Ministério Público, CNJ: Conselho Nacional de Justiça.

3.3 Nível Micro: Desempenho Laboratorial

O terceiro nível concentra-se nos indicadores consolidados pelo projeto FORESIGHT (Houck & Speaker, 2020; 2024; Speaker, 2021). Nesse nível, propõe-se a divisão dos indicadores em três categorias: operacionais, de qualidade técnica e de recursos.

Dentre os indicadores operacionais, o TAT (tempo entre a requisição e a expedição do laudo) constitui o indicador primário de celeridade. A mediana oferece tendência central imune a outliers; o P90 captura o efeito de "cauda longa". A razão P90/Mediana fornece medida de dispersão: valores próximos a 1 indicam distribuição homogênea; valores superiores a 3 indicam variabilidade problemática. A decomposição do TAT (Omari *et al.*, 2018), permite identificar gargalos específicos. O TAT mediano não distingue complexidade de casos e pode ser influenciado por devoluções administrativas. O FORESIGHT situa a mediana entre 30 e 46 dias e o P90 em 90 dias. O número de Casos por FTE por Ano avalia produtividade individual. O FORESIGHT sugere 95 casos/FTE/ano, valor que mascara variação entre disciplinas forenses e portes de laboratório. As limitações incluem não ponderação por complexidade e não comparabilidade entre espécies de exame. O backlog superior a 30 dias indica proporção de casos pendentes além deste limiar. O FORESIGHT sugere 15% como aceitável. Valores elevados indicam desbalanceamento entre demanda e capacidade, que tende a agravar-se exponencialmente. A Taxa de Conclusão mede requisições finalizadas em relação ao total recebido. A Taxa de Devolução avalia requisições devolvidas sem laudo, podendo decorrer de motivos legítimos ou problemas de interface.

Para qualidade técnica, sugere-se a Taxa de Identificação Positiva, que mede a proporção de exames com identificação de pessoa ou objeto. A literatura não estabelece uma taxa "ideal". Taxas baixas podem indicar problemas técnicos ou qualidade do material. Por outro lado, taxas elevadas podem indicar excelência técnica ou viés confirmatório. As Taxas de Falso Positivo (TFP) e Falso Negativo (TFN) indicam, respectivamente, identificações incorretas e oportunidades perdidas. Em casos reais, nos quais a verdade fundamental é desconhecida, essas taxas só podem ser estimadas via testes de proficiência. Ulery e colaboradores (2011) estimaram TFP em 0,1% e TFN em 7,5%. O Coeficiente Kappa de Cohen mede concordância inter-examinadores, sendo $k = (Po - Pe)/(1 - Pe)$, onde Po representa a proporção

de acordo observado, e Pe a proporção esperada ao acaso. Valores de k acima de 0,80 indicam convergência metodológica, enquanto valores inferiores sugerem necessidade de padronização ou treinamento (Landis & Koch, 1977).

Como indicadores de recursos, sugere-se a Razão Perito/Demanda, que equaciona efetivo disponível com volume de requisições. O número de Horas por Caso indica dispêndio técnico para cada exame. Ambas as medidas, fracionadas por espécie de exame, permitem isolar variação intraespécies. O Investimento em Capacitação, medido em horas de treinamento por perito por ano, avalia compromisso organizacional com desenvolvimento de capital humano. A meta sugerida, derivada de requisitos ASCLD/LAB e ISO 17025, situa-se acima de 40 horas/perito/ano. A Tabela 5 sintetiza os indicadores de nível meso distinguindo as categorias, o método de cálculo e os valores de referência encontrados na literatura.

Tabela 5 – Resumo dos indicadores do nível micro para o desempenho laboratorial.

Categoria	Indicador	Forma de calcular	Valor de referência
Operacionais	TAT mediano	Mediana (dias)	30-46 dias
	TAT P90	Percentil 90	90 dias
	Razão P90/Mediana	$P90 \div \text{Mediana}$	< 3
	Casos/FTE/ano	$\text{Casos} \div \text{FTE}$	95
	Backlog >30 dias	% pendentes >30d	< 15%
	Taxa conclusão	$\text{Expedidos} \div \text{Criados}$	> 85%
Qualidade	Taxa identificação positiva	% identificações	Sem meta (interpretar com cautela)
	Taxa falso positivo	Via proficiência	0,1% (Ulery et al., 2011)
	Taxa falso negativo	Via proficiência	7,5% (Ulery et al., 2011)
	Kappa Cohen	$(Po-Pe)/(1-Pe)$	> 0,80
Recursos	Perito/Demanda	$\text{FTE} \div \text{Requisições}$	Deficit EUA: 19,7%
	Horas/caso	$\text{H técnicas} \div \text{Casos}$	Por espécie de exame
	Capacitação	$\text{H treinamento/servidor/ano}$	> 40h

Fonte: Adaptado de Houck & Speaker (2020; 2024), Speaker (2021) e Ulery et al. (2011). Notas: TAT: *turnaround time*, FTE: *full-time equivalente*.

3.4 Proposta Metodológica do Nível de Integração

Apesar de os índices de síntese apresentarem limitações, o seu acompanhamento permite uma avaliação baseada em evidência e possibilita, em potencial, uma atuação gerencial mais assertiva. A literatura forense é omissa no estabelecimento de relações quantitativas entre a eficiência laboratorial e resultados do SJC. Peterson et al. (2013), analisando 4.205 casos em cinco jurisdições estadunidenses, apontou que a evidência

forense coletada constitui um preditor de prisão com razão de chances de aproximadamente 1,66. Entretanto, essa razão de chances apresentou efeito dependente do tempo de processamento da amostra forense e não foi estatisticamente superior a evidências não-forenses. Entretanto, Baskin & Sommers (2012), analisando 1.940 casos de agressão e roubo nas mesmas jurisdições, encontraram que variáveis de evidência forense não foram preditoras estatisticamente significativas de prisão, acusação ou condenação quando controladas por outros fatores – sendo relatos de vítimas e testemunhas os principais determinantes da progressão processual. Essa contradição metodológica foi reconhecida pela revisão do Home Office (2022), que indica que “poucos estudos capturaram de forma robusta e abrangente o valor que a ciência forense agrega ao” SJC.

Mais especificamente, não foram encontrados estudos que correlacionem indicadores criminológicos e métricas laboratoriais. A prática e a doutrina assumem que “justiça tardia é justiça negada”, entretanto os impactos da eficiência laboratorial – ou mesmo a relação causal entre redução de TAT e taxa de elucidação – não são mensurados. Essa ausência de evidências impossibilita validação de índices compostos. Não obstante, a experiência brasileira com índices multidimensionais aponta que, mesmo ante limitações de validação, instrumentos metodologicamente fundamentados podem: (i) organizar informações dispersas em formato acessível a gestores; (ii) orientar priorização de recursos; (iii) viabilizar comparações temporais e inter-institucionais; (iv) estimular acumulação gradual de evidências através de aplicação sistemática (Costa *et al.*, 2018).

Isto posto, propõe-se o Índice de Desempenho Laboratorial Forense (IDLF). Embora o adágio “o que é medido é gerenciado” seja frequentemente invocado para justificar sistemas de indicadores, sua origem remonta ao alerta de Ridgway (1956) sobre consequências disfuncionais da mensuração. Esse reconhecimento fundamenta a proposta de que o IDLF constitui ferramenta diagnóstica, não meta em si mesma. O índice visa oferecer uma ferramenta de síntese para a gestão laboratorial e, ao mesmo tempo, acumular a base empírica necessária para sua própria evolução. Sua proposta prioriza a metodologia, a capacidade de decomposição analítica e o reconhecimento transparente de suas limitações.

O IDLF agrega indicadores dos três níveis (macro, meso, micro) em métrica única, que varia entre 0,0 (desempenho ótimo) e 1,0 (desempenho crítico), onde valores mais elevados indicam a necessidade de intervenção da forma

$$IDLF = \sum w_k D_k = w_{macro} \cdot D_{macro} + w_{meso} \cdot D_{meso} + w_{micro} \cdot D_{micro} \quad (1)$$

onde D_k representa a dimensão normalizada do nível k , e w_k são pesos que satisfazem

$$\sum w_k = 1. \quad (2)$$

Os pesos sugeridos são $w_{macro} = 0,25$, $w_{meso} = 0,25$ e $w_{micro} = 0,50$, para preponderância de indicadores diretos do desempenho laboratorial, reconhecendo relevância do contexto (macro) e da integração institucional (meso). Cada dimensão D_k constitui média ponderada dos indicadores normalizados daquele nível

$$D_k = \sum_{i=1}^{N_k} w_{k,i} I_{k,i,norm}, \quad (3)$$

onde $I_{k,i,norm}$ é o indicador i do nível k após normalização, e $w_{k,i}$ são pesos específicos que satisfazem, para cada nível,

$$\sum_{i=1}^{N_k} w_{k,i} = 1. \quad (4)$$

A normalização converte indicadores com escalas distintas para intervalo comum $[0, 1]$, viabilizando agregação matematicamente válida com transformação min-max (Costa *et al.*, 2018). Para os indicadores decrescentes – melhores quando menores – e.g., TAT, backlog, taxa de falso positivo –, a normalização é da forma

$$I_{norm} = \frac{I - I_{min}}{I_{max} - I_{min}}, \quad (5)$$

onde I representa o valor do indicador e I_{min} e I_{max} , respectivamente, seus valores mínimo e máximo. Para os indicadores crescentes – melhores quando maiores – e.g., produtividade, taxa de verdadeiro positivo –, a normalização é da forma

$$I_{norm} = \frac{I_{max} - I}{I_{max} - I_{min}}. \quad (6)$$

O valor mínimo do índice pode ser assumido como a melhor situação empiricamente observável. Para TAT e backlog, zero representa valor mínimo teórico. Para taxas percentuais em que zero é inatingível, o valor mínimo pode ser estabelecido como percentil 5 da distribuição da série histórica ou por meio de referências da literatura internacional. Para o valor máximo, sugere-se

a média mais dois desvios-padrão da distribuição da série histórica limitado ao pior resultado observado.

A atribuição de pesos busca agregar os indicadores respeitando as particularidades para evitar a preponderância de um indicador sobre os demais (Costa *et al.*, 2018, p. 25). Essa escolha metodológica reconhece a ausência de evidências empíricas que fundamentem ponderação diferencial *a priori*. Porém, estudos de validação poderiam identificar capacidades preditivas diferenciadas dos indicadores.

Dessa forma, propõe-se a distribuição uniforme dos pesos entre categorias dentro de cada nível, preservando a equivalência entre as dimensões conceituais. No nível macro, os dez indicadores distribuídos em três categorias (Tabela 3) recebem pesos por categoria de $w_{macro,cat} = 0,25/3$, distribuídos internamente com quatro indicadores estruturais com $w_{macro,estrut,i} = (0,25/3)/4$; dois indicadores de criminalidade com $w_{macro,crim,i} = (0,25/3)/2$; e quatro indicadores sistêmicos com $w_{macro,sist,i} = (0,25/3)/4$.

No nível meso, os nove indicadores distribuídos em três categorias (Tabela 4) recebem pesos por categoria de $w_{meso,cat} = 0,25/3$, distribuídos como: quatro indicadores de integração institucional com $w_{meso,integ,i} = (0,25/3)/4$; três indicadores de capacidade do SJC com $w_{meso,cap,i} = (0,25/3)/3$; e dois indicadores de qualidade processual com $w_{meso,qual,i} = (0,25/3)/2$.

No nível micro, os 13 indicadores distribuídos em três categorias (Tabela 5) recebem pesos por categoria de $w_{micro,cat} = 0,50/3$, distribuídos como: seis indicadores operacionais com $w_{micro,oper,i} = (0,50/3)/6$; quatro de qualidade técnica com $w_{micro,qual,i} = (0,50/3)/4$; e três de recursos com $w_{micro,rec,i} = (0,50/3)/3$. Sugere-se ainda a adoção de cinco faixas de classificação que orientam a interpretação do índice e facilitam a comunicação, conforme indicado na Tabela 6. Naturalmente, essa distribuição de pesos e faixas necessita de revisão futura baseada em análise fatorial que teste se o agrupamento conceitual corresponde à estrutura empírica latente.

Outra contribuição metodológica reside na capacidade de decomposição do IDLF por espécie de exame, permitindo diagnósticos granulares sobre desempenho setorial do laboratório. Laboratórios forenses são organizações técnicas heterogêneas onde diferentes setores (papiloscopia, DNA, balística, química forense etc.) operam com: (i) tecnologias distintas; (ii) perfis de complexidade diferenciados; (iii) pressões de demanda variáveis; e (iv) graus diversos de padronização metodológica. Agregar todos os exames em único índice pode mascarar desempenhos setoriais divergentes.

Tabela 6 – Interpretação do valor do IDLF.

Faixa IDLF	Classificação	Interpretação Gerencial
0,000 - 0,200	Desempenho muito bom	Laboratório próximo a avaliações internacionais; priorizar a manutenção
0,201 - 0,300	Desempenho bom	Acima da média nacional; oportunidades pontuais de melhoria
0,301 - 0,400	Desempenho médio	Alinhado com média nacional; requer atenção a gargalos específicos
0,401 - 0,500	Desempenho insuficiente	Abaixo da média; necessita intervenções estruturadas
0,501 - 1,000	Desempenho crítico	Défices severos; demanda investimento prioritário

Fonte: elaborado pelo autor.

O IDLF por espécie de exame ($IDLF_e$) utiliza mesma estrutura metodológica do IDLF geral, mas calculado exclusivamente com dados daquela espécie, sendo

$$IDLF_e = \sum w_k D_k = w_{macro} \cdot D_{macro} + w_{meso} \cdot D_{meso,e} + w_{micro} \cdot D_{micro,e} \quad (7)$$

onde D_{macro} permanece invariante (contexto sistêmico não varia por espécie), enquanto $D_{meso,e}$ e $D_{micro,e}$ são calculados com dados específicos da espécie quando disponíveis. No exemplo de resultados da Tabela 7, o laboratório apresenta desempenho médio global ($IDLF = 0,333$), mas a análise desagregada revela crise específica no setor de DNA ($IDLF = 0,428$) que demanda atenção prioritária. Sem decomposição, o índice médio mascararia essa situação. É importante salientar que o IDLF global, como proposto, não dispensa a necessidade de atuação gerencial, pois a análise das dimensões e as espécies desagregadas permitem a identificação de gargalos e tomadas de decisões menos suscetíveis a erros.

Tabela 7 – Exemplo de decomposição do IDLF geral por espécies de exame.

Espécie	D_{macro}	D_{meso}	D_{micro}	$IDLF_e$
Papiloscopia	0,350	0,280	0,220	0,268 (bom)
DNA	0,350	0,320	0,520	0,428 (insuficiente)
Balística	0,350	0,290	0,310	0,315 (médio)
IDLF Geral	0,350	0,297	0,350	0,333

Fonte: elaborado pelo autor.

4 DISCUSSÃO

A presente seção visa discutir as limitações práticas e tecnológicas na

proposta do Índice de Desempenho Laboratorial Forense (IDLF), as principais contribuições teóricas e práticas, bem como as limitações. Em complementação, propõem-se estratégias de superação fundamentadas em tecnologias emergentes e arranjos institucionais colaborativos.

A primeira limitação é a fragmentação dos dados que demanda, pelo menos, as seguintes ações: a adoção de uma chave única relacionada à *notitia criminis* que permeie o SJC, passando por requisição pericial, laudo, inquérito, denúncia e sentença; a interoperabilidade dos sistemas interagências, respeitando requisitos de segurança e legislação; e arquitetura em que cada instituição mantém autonomia sobre seus dados, disponibilizando indicadores agregados. Como potencial, o Sistema Nacional de Informações de Segurança Pública (Sinesp) oferece arcabouço potencial, embora sua cobertura permaneça incompleta (FBSP, 2023).

Por outro lado, a complexidade e o volume de dados envolvidos na construção do IDLF justificam o investimento em técnicas de *big data* e inteligência artificial. Entre as técnicas sugeridas está a extração automatizada de informações de documentos (laudos, inquéritos, processos, etc...) não estruturados utilizando técnicas de múltiplos agentes de processamento de linguagem natural. Esses agentes podem rodar nos níveis micro, do laboratório forense, e nos níveis meso, consolidando dados da polícia judiciária e dos processos tramitados. O modelo multiagente distribui responsabilidade de coleta entre instituições, permite a escalabilidade vertical e horizontal dos índices, podendo ser agregados a outras métricas de qualidade e possibilita a validação cruzada.

Outra dificuldade encontra-se na necessidade da implantação de programas de proficiência e de verificação de qualidade por reexame às cegas. Indicadores de qualidade técnica (taxa de falso positivo, concordância inter-peritos) dependem de verificação de proficiências periódicas. Um exemplo de verificações é o Exercício Colaborativo Internacional do programa do Escritório das Nações Unidas sobre Drogas e Crimes (UNODC), que avalia a qualidade da capacidade analítica de um laboratório. Existem também empresas acreditadas por órgão internacionais que aplicam verificação de proficiência. Por fim, tem-se o desafio de promover a capacitação constante dos profissionais, tanto técnica como gerencial.

Outro ponto de crítica é que a proposta ainda perpassa por muitos fatores conceituais, e a ausência de muitos dados limita uma validação empírica e um ajuste metodológico. Junta-se a esse fato a escassez de dados padronizados, que persiste desde o diagnóstico da Senasp (2013). Muitos

indicadores propostos, especialmente no nível meso, são teoricamente utilizáveis, mas praticamente inacessíveis nas condições atuais de fragmentação informacional do SJC brasileiro.

Frente às limitações, faz-se importante pontuar que o desempenho laboratorial vai além da eficiência operacional sob responsabilidade do gestor laboratorial (nível micro), mas também depende de fatores criminológicos exógenos (nível macro) e da utilização do produto pericial, que depende de atores externos (nível meso). Dessa forma, a melhoria dos serviços prestados, incluindo as análises forenses, implica responsabilidade compartilhada entre múltiplos atores do Sistema de Justiça Criminal. Na prática, os índices que compõem o IDLF podem ser interpretados como ferramenta diagnóstica que permite identificar onde residem os gargalos – se no contexto, na interface institucional ou na operação interna –, orientando intervenções para o ponto de maior alavancagem. Entretanto, a arquitetura comum viabiliza a construção de futuros índices, como, por exemplo, um índice de desempenho do sistema de justiça criminal.

Como contribuição, o presente trabalho busca uma articulação entre as perspectivas gerencialistas e criminológicas em estrutura analítica. O projeto FORESIGHT (Houck & Speaker, 2020; 2024) focaliza indicadores operacionais assumindo que eficiência interna traduz-se em efetividade. Por outro lado, a crítica criminológica (Vargas, 2014; Ribeiro & Diniz, 2020) aponta funis entre etapas que comprometem essa pressuposição, mas não propõe instrumentos gerenciais alternativos. Outra contribuição é que a limitação de mensuração e proposição de estratégias para realizá-las é fundamentada em tecnologias emergentes (interoperabilidade, *big data*, sistemas multiagente) e arranjos institucionais colaborativos (programas de proficiência, responsabilidade compartilhada).

5 CONSIDERAÇÕES FINAIS

O presente trabalho propôs uma estrutura de avaliação multinível para laboratórios forenses que concilia a necessidade de gestão eficiente – materializada em indicadores operacionais consolidados internacionalmente – com o reconhecimento de que a demanda por serviços periciais deriva de dinâmicas criminológicas multifatoriais exógenas ao próprio laboratório. A estrutura organiza indicadores em quatro níveis de análise (macro, meso, micro, integração), oferecendo especificação metodológica para síntese em índice único (IDLF - Índice de Desempenho Laboratorial Forense).

A contribuição central reside na mudança de perspectiva em que o laboratório forense atua como componente de um sistema complexo, cujo

funcionamento depende de múltiplos atores com responsabilidades compartilhadas. Essa perspectiva contextualiza resultados operacionais em ambiente criminológico e institucional, além de identificar se há gargalos no contexto, na interface ou na operação.

O IDLF proposto oferece ferramenta gerencial que visa sintetizar informações dispersas em formato acessível a gestores, dentro de suas próprias limitações. Para gestores, recomenda-se priorizar a consolidação de infraestrutura informacional com registros desagregados por espécie de exame, iniciar coleta sistemática de indicadores micro e articular mecanismos de troca de informações com Polícia Civil, Ministério Público e Judiciário.

Como propostas de continuidade, destaca-se, no eixo tecnológico, avanços em inteligência artificial e *big data* para automatizar processos de coleta, normalização e análise que atualmente exigiriam esforço manual proibitivo. O desafio de avaliar desempenho de laboratórios forenses ilustra uma relação de compromisso entre a eficiência gerencial na prestação do serviço público e a complexidade sistêmicas da defesa social. A proposta apresentada pretende oferecer instrumento de quantificação, reconhecendo que o laboratório forense efetivo não é apenas o que apresenta indicadores operacionais favoráveis, mas o que contribui – em articulação com a polícia judiciária e demais agentes do SJC – para que a justiça seja feita. O Autor gostaria de agradecer ao apoio da Rede Mineira de Ciências Forenses e a Fapemig, por meio do RED-00120-23.

REFERÊNCIAS

ALFARO, C.; MADRIGAL, G.; HERNÁNDEZ, M. Improving forensic processes performance: A Lean Six Sigma approach. **Forensic Science International: Synergy**, v. 2, p. 90-94, 2020.

BANCO MUNDIAL. **Confronting the violence epidemic: A policy note on crime and violence in Brazil**. Washington: World Bank, 2018.

BASKIN, D.; SOMMERS, I. The influence of forensic evidence on the case outcomes of assault and robbery incidents. **Criminal Justice Policy Review**, v. 23, n. 2, p. 186-210, 2012.

BEATO FILHO, C. **Crime e cidades**. Belo Horizonte: Editora UFMG, 2012.

BENSON, W.; STACY, J.; WORLEY, M. **Systems Analysis of Criminalistics Operations**. Washington, DC: Law Enforcement Assistance Administration (LEAA), U.S. Department of Justice, 1970.

BORG, M. J.; PARKER, K. F. Mobilizing law in urban areas: The social structure of

homicide clearance rates. **Law & Society Review**, v. 35, n. 2, p. 435-466, 2001.

BRESSER-PEREIRA, L. C. Da administração pública burocrática à gerencial. **Revista do Serviço Público**, v. 47, n. 1, p. 7-40, 1996.

CANO, I.; DUARTE, T. L. A mensuração da impunidade no sistema de justiça criminal do Rio de Janeiro. **Segurança, Justiça e Cidadania**, v. 4, p. 9-44, 2010.

CARDIA, N.; ADORNO, S.; POLETO, F. Homicídios e violação de direitos humanos em São Paulo. **Estudos Avançados**, v. 17, n. 47, p. 43-73, 2003.

CERQUEIRA, D. **Mapa dos homicídios ocultos no Brasil**. Texto para Discussão IPEA, n. 1848. Brasília: IPEA, 2013.

CERQUEIRA, D.; LOBÃO, W. Determinantes da criminalidade: arcabouços teóricos e resultados empíricos. **Dados**, v. 47, n. 2, p. 233-269, 2004.

CERQUEIRA, D. **Causas e consequências do crime no Brasil**. Brasília: Banco Nacional de Desenvolvimento Econômico e Social, 2014.

CNJ. **CNJ lança o Mapa Nacional do Júri e solicita levantamento aos tribunais**. Portal CNJ, 2024. Disponível em: <https://www.cnj.jus.br/cnj-lanca-o-mapa-nacional-do-juri-e-solicita-levantamento-aos-tribunais/>. Acesso em: dez. 2025.

CNJ. **CNJ amplia critérios de prioridade de julgamento para o Mês Nacional do Júri**. Portal CNJ, 2025. Disponível em: <https://www.cnj.jus.br/cnj-amplia-criterios-de-prioridade-de-julgamentos-para-o-mes-nacional-do-juri/>. Acesso em: jan. 2026.

COHEN, L.; FELSON, M. Social change and crime rate trends: A routine activity approach. **American Sociological Review**, v. 44, n. 4, p. 588-608, 1979.

COOK, P. J. The clearance rate as a measure of criminal justice system effectiveness. **Journal of Public Economics**, v. 11, n. 1, p. 135-142, 1979.

COSTA, A. A (in)efetividade da justiça criminal brasileira: uma análise do fluxo de justiça dos homicídios no Distrito Federal. **Civitas**, v. 15, n. 1, p. 11-26, 2015.

COSTA, M. A. *et al.* **Atlas da Vulnerabilidade Social no Brasil**: IVS. Brasília: IPEA, 2018.

COSTA, M.; MARGUTI, B. (Ed.). **Atlas da vulnerabilidade social nos municípios brasileiros**. Brasília: IPEA, 2015.

ETERNO, J. A.; SILVERMAN, E. B. **The crime numbers game**: Management by manipulation. Boca Raton: CRC Press, 2012.

FAJNZYLBER, P.; LEDERMAN, D.; LOAYZA, N. Inequality and violent crime. **Journal of Law and Economics**, v. 45, n. 1, p. 1-40, 2002.

FBSP. **19º Anuário Brasileiro de Segurança Pública 2025**. São Paulo: Fórum

Brasileiro de Segurança Pública, 2025. Disponível em: <https://publicacoes.forumseguranca.org.br/items/c3605778-37b3-4ad6-8239-94e4cb236444> . Acesso em: mar. 2026.

GOODHART, C. Problems of monetary management: The UK experience. In: GOODHART, C. **Monetary Theory and Practice**. London: Macmillan, 1984. p. 91-121.

HOME OFFICE. **Impact of Forensic Science Project**: Phase 1 Report. London: Home Office, 2022. Disponível em: <https://www.gov.uk/government/publications/impact-of-forensic-science-project>. Acesso em: dez. 2025.

HOOD, C. A public management for all seasons? **Public Administration**, v. 69, n. 1, p. 3-19, 1991.

HOOD, C. The "new public management" in the 1980s: Variations on a theme. **Accounting, Organizations and Society**, v. 20, n. 2-3, p. 93-109, 1995.

HOUCK, M.; SPEAKER, P. Project FORESIGHT: A ten-year retrospective. **Forensic Science International: Synergy**, v. 2, p. 275-281, 2020. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2589871X20300541>. Acesso em: dez. 2025.

HOUCK, M.; SPEAKER, P. **Project FORESIGHT Annual Report, 2020-2021**. Morgantown: West Virginia University, 2022.

HOUCK, M.; SPEAKER, P. **Forensic Science Laboratory Benchmarking**: The FORESIGHT Manual. Boca Raton: CRC Press, 2024. Disponível em: <https://doi.org/10.4324/9780429318528>. Acesso em: dez. 2025.

IPEA. Atlas da Violência 2016. Brasília: IPEA, 2016.

KLEINERT, A. Der messende Luchs: Zwei verbreitete Fehler in der Galilei-Literatur. NTM Zeitschrift für Geschichte der Wissenschaften, **Technik und Medizin**, v. 17, n. 2, p. 199-206, 2009.

LANDIS, J.; KOCH, G. The measurement of observer agreement for categorical data. **Biometrics**, v. 33, n. 1, p. 159-174, 1977.

LIMA, R.; RATTON, J.; AZEVEDO, R. (Org.). **Crime, polícia e justiça no Brasil**. São Paulo: Contexto, 2014.

MANDL, U.; DIERX, A.; ILZKOVITZ, F. The effectiveness and efficiency of public spending. **European Economy Economic Papers**, n. 301. Brussels: European Commission, 2008.

MANNING, P. K. Theorizing policing: The drama and myth of crime control in the NYPD. **Theoretical Criminology**, v. 5, n. 3, p. 315-344, 2001.

MARTIN, T. **Galilée**: les droits de la science et la méthode des sciences

physiques. Paris: Didier, 1868.

MEIER, K. J.; O'TOOLE JR., L. J.; RUTGERS, M. R. Effectiveness, efficiency, and equity tradeoffs in public programs: A citizen experiment. **Public Administration Review**, v. 83, n. 4, p. 829-844, 2023.

MINAS GERAIS. Polícia Civil. Diretoria de Estatística e Análise Criminal. **Relatório estatístico: Mortes Violentas Intencionais (MVI) e Indicador de Elucidação**: art. 4º-A, da Lei Estadual nº 13.772/2000 (2º semestre de 2023). Belo Horizonte: DEACrim, 2024. Disponível em: https://www.seguranca.mg.gov.br/images/0_planilhas-e-pdfs/transparencia/documentos/Anexo_93839031_RELATORIO__Art._4_C2_BA_A_da_Lei_n_C2_BA_13772_2000__2_C2_BA_Semestre_de_2023__281_29_assinado_assinado.pdf. Acesso em: 30 mar. 2026.

OFFICE OF LAW ENFORCEMENT ASSISTANCE (OLEA). **Crime Laboratories: Three Study Reports**. Final Report to Office of Law Enforcement Assistance. Washington, DC: U.S. Department of Justice, 1968.

OMARI, S. *et al.* Turnaround time decomposition in a forensic laboratory: Analysis and improvement opportunities. **Forensic Science International**, v. 292, p. 1-8, 2018.

PARKER, B. The status of forensic science in the administration of criminal justice. **Revista Jurídica de la Universidad de Puerto Rico**, v. 32, p. 405-419, 1963.

PETERSON, J. *et al.* **The role and impact of forensic evidence in the criminal justice process**. Washington: NIJ, 2013.

POLLITT, C. **Managerialism and the public services: The Anglo-American experience**. Oxford: Blackwell, 1990.

RIBEIRO, L.; DINIZ, A. M. A. The flow of murder cases through the criminal justice system in a Brazilian city. **Homicide Studies**, v. 24, n. 4, p. 331-353, 2020.

RIBEIRO, L.; SILVA, K. Fluxo do Sistema de Justiça Criminal Brasileiro: Um balanço da literatura. **Cadernos de Segurança Pública**, ano 2, n. 1, p. 15-27, 2010. Disponível em: <http://www.isprevista.rj.gov.br/download/Rev20100102.pdf>. Acesso em: dez. 2025.

RIDGWAY, V. F. Dysfunctional consequences of performance measurements. **Administrative Science Quarterly**, v. 1, n. 2, p. 240-247, 1956.

RUTGERS, M.; VAN DER MEER, H. The origins and restriction of efficiency in public administration: Regaining efficiency as the core value of public administration. **Administration & Society**, v. 42, n. 7, p. 755-779, 2010.

SAPORI, L. **Segurança pública no Brasil: desafios e perspectivas**. Rio de Janeiro: Editora FGV, 2007.

SEN, A. K. **Desenvolvimento como Liberdade**. São Paulo: Companhia das

Letras, 2000.

SENASP. **Diagnóstico da Perícia Criminal**. Brasília: Secretaria Nacional de Segurança Pública, 2013.

SPEAKER, P. J. The Forensic Laboratory Workforce Calculator. **Forensic Science Policy & Management**, v. 12, n. 1, p. 1-10, 2021.

ULERY, B. *et al.* Accuracy and reliability of forensic latent fingerprint decisions. **Proceedings of the National Academy of Sciences**, v. 108, n. 19, p. 7733-7738, 2011.

UNODC. **Global Study on Homicide 2019**. Vienna: United Nations Office on Drugs and Crime, 2019.

VARGAS, J. D. **Estupro**: que justiça? Fluxo do funcionamento e análise do tempo da justiça criminal para o crime de estupro. 2004. Tese (Doutorado) – IUPERJ, Rio de Janeiro, 2004.

VARGAS, J. Fluxo do sistema de justiça criminal. In: LIMA, R.; RATTON, J.; AZEVEDO, R. (Org.). **Crime, polícia e justiça no Brasil**. São Paulo: Contexto, 2014. p. 412-423.

WEISBURD, D., MASTROFSKI, S., MCNALLY, A., GREENSPAN, R., WILLIS, J. Reforming to preserve: Compstat and strategic problem solving in American policing. **Criminology & Public Policy**, v. 2, n. 3, p. 421-456, 2003.