

ROBÔ-GLOSA: AUDITORIA AUTOMATIZADA DE BAIXO CUSTO BASEADA EM REGRAS EM UMA OPERADORA DE AUTOGESTÃO BRASILEIRA

Robô-Glosa: Low-Cost Rule-Based Automated Auditing in a Brazilian Self-Managed Health Insurance Provider

AUTORES

Guilherme Coelho

1 Universidade Estadual de
Campinas

CONTATO

Guilherme Coelho

1 guilcoel@unicamp.br

Obs: este artigo reflete exclusivamente a opinião do(s) autor(es), não representando a posição do IESS.

RESUMO

Introdução: A auditoria de contas médicas é fundamental para a sustentabilidade financeira do sistema de saúde suplementar brasileiro. Com mais de 53 milhões de beneficiários, as operadoras de autogestão enfrentam desafios particulares de volume e de padronização. **Objetivo:** Descrever o desenvolvimento e a validação preliminar do Robô-Glosa, um sistema automatizado de auditoria de contas médicas baseado em regras, implementado em uma operadora brasileira de autogestão. **Métodos:** Foi desenvolvido um sistema em Visual Basic for Applications (VBA v5.7), parametrizável, com 11 regras de conformidade clínica, contratual e estatística. A unidade de análise foi a autorização (handle). O desempenho foi comparado ao da auditoria manual conduzida por enfermeira auditora, utilizando-se acurácia, sensibilidade, especificidade, valor preditivo positivo (VPP) e valor preditivo negativo (VPN), com intervalos de confiança de 95% (método de Wilson). **Resultados:** Foram analisadas 1.909 autorizações únicas referentes a uma semana de faturamento hospitalar (maio de 2025). O sistema alcançou acurácia de 94,71% (IC 95%: 93,61–95,63), sensibilidade de 93,40% (IC 95%: 92,05–94,54), especificidade de 100% (IC 95%: 98,99–100), VPP de 100% (IC 95%: 99,73–100) e VPN de 78,91% (IC 95%: 75,04–82,33), com zero falsos positivos. O tempo de execução foi inferior a 1 segundo, em contraste com aproximadamente 2 dias de trabalho manual. Foram identificados 101 falsos negativos, correspondentes a glosas que dependiam de julgamento clínico contextual. **Conclusão:** O Robô-Glosa demonstrou viabilidade como ferramenta de pré-triagem de baixo custo, com alto valor preditivo positivo e potencial para reduzir a carga operacional da auditoria em saúde suplementar.

Palavras-chave: Administração em Saúde; Auditoria Médica; Automação; Eficiência

ABSTRACT

Introduction: Medical claims auditing is critical to the financial sustainability of the Brazilian supplementary healthcare system. With over 53 million beneficiaries, self-managed health plans face particular challenges regarding audit volume and process standardization. **Objective:** To describe the development and preliminary validation of Robô-Glosa, a rule-based automated medical claims auditing system implemented in a Brazilian self-managed health insurance provider. **Methods:** A parameterizable system was developed in Visual Basic for Applications (VBA v5.7) that incorporates 11 clinical, contractual, and statistical compliance rules. The unit of analysis was the authorization (handle). Performance was compared against manual auditing conducted by a nurse auditor, using accuracy, sensitivity, specificity, and precision with 95% confidence intervals. **Results:** A total of 1,909 unique authorizations from one week of hospital billing data (May 2025) were analyzed. The system achieved accuracy of 94.71% (95% CI: 93.71–95.71), sensitivity of 93.40% (95% CI: 92.16–94.65), specificity of 100% (95% CI: 99.21–100), and precision of 100% (95% CI: 99.79–100), with zero false positives. Execution time was under 1 second, compared to approximately 2 days of manual review. A total of 101 false negatives were identified, corresponding to denials requiring contextual clinical judgment. **Conclusion:** Robô-Glosa demonstrated feasibility as a low-cost pre-screening tool with high precision and potential to reduce the operational burden of supplementary health auditing.

Keywords: Automation; Health Administration; Health Information Systems; Medical Audit; Supplementary Health

INTRODUÇÃO

A auditoria de contas médicas é um dos pilares da gestão financeira e assistencial no sistema de saúde suplementar brasileiro. Em outubro de 2025, o setor contava com 53,3 milhões de vínculos de beneficiários em planos de assistência médica (BRASIL, 2025a), e a Agência Nacional de Saúde Suplementar (ANS) tem reforçado progressivamente as exigências de transparência, qualidade e controle nos serviços prestados pelas operadoras (BRASIL, 2024). Nesse cenário, a auditoria não se restringe à conferência de valores: constitui instrumento de detecção de inconformidades, de educação permanente e de fomento à eficiência operacional (GAMARRA, 2018; PINHEIRO et al., 2016).

As operadoras de autogestão ocupam uma posição singular nesse ecossistema. Segundo dados da UNIDAS, a modalidade reúne aproximadamente 3 milhões de beneficiários, com base populacional mais envelhecida do que a média do setor e despesas assistenciais anuais da ordem de R\$ 21,5 bilhões (UNIDAS, 2025). Suas estruturas administrativas são tipicamente enxutas, com equipes de auditoria reduzidas e recursos tecnológicos limitados. A Resolução Normativa nº 649/2025 da ANS reforçou a necessidade de governança e sustentabilidade nessa modalidade (BRASIL, 2025b), tornando ainda mais premente a busca por processos internos eficientes e padronizados.

Apesar de seu valor estratégico, a realização de auditorias manuais enfrenta barreiras conhecidas: sobrecarga de trabalho, variabilidade subjetiva na interpretação de regras e limitações dos sistemas de informação (HANSKAMP-SEBREGTS et al., 2019; SCARPARO et al., 2009). No contexto brasileiro, a revisão linha a linha de planilhas de faturamento hospitalar, ainda predominante em operadoras de menor porte, consome dias de trabalho por semana e está sujeita a erros operacionais e inconsistências entre auditores (CRISTINA et al., 2021; ARAYA; PINHEIRO; BEZERRA, 2016; BAURAKIADES et al., 2012).

Esses desafios abrem espaço para a automação. Embora a literatura internacional registre avanços no uso de inteligência artificial e de análise preditiva aplicadas à auditoria em saúde (RULE; CHIANG; HRIBAR, 2020), a maior parte das soluções descritas depende de plataformas proprietárias ou de infraestrutura computacional de alto custo, o que limita sua adoção por operadoras com orçamentos restritos. No Brasil, experiências recentes com métodos quantitativos acessíveis para classificação de risco em prestadores de saúde sugerem que abordagens simples e de baixo custo podem ter impacto relevante na prática da auditoria (COELHO, 2025). Até onde se tem

conhecimento, não há relatos publicados de sistemas automatizados de auditoria baseados em planilha eletrônica, com validação formal de desempenho, no contexto da saúde suplementar brasileira.

Este relato de experiência tem como objetivo descrever o desenvolvimento e a validação preliminar do Robô-Glosa, uma ferramenta automatizada de auditoria de contas médicas baseada em 11 regras de conformidade clínica, contratual e estatística, desenvolvida em VBA/Excel e implementada em uma operadora brasileira de autogestão.

MÉTODOS

Tipo de estudo, local e período

Trata-se de um relato de experiência sobre avaliação de desempenho, que descreve o desenvolvimento e a validação preliminar de um sistema automatizado de auditoria de contas médicas em uma operadora de autogestão em saúde suplementar localizada na região Sudeste do Brasil. O desenvolvimento do Robô-Glosa teve início em maio de 2025, com ciclos iterativos de programação, teste e refinamento das regras junto à equipe de auditoria, alcançando a versão 5.7 utilizada neste estudo. Os dados analisados para validação referem-se ao período de 26 a 29 de maio de 2025 e abrangem o faturamento hospitalar dos prestadores da rede credenciada.

Equipe envolvida

A implementação envolveu quatro profissionais: um médico auditor (desenvolvedor principal, responsável pela programação em VBA, definição dos critérios clínicos e análise de dados), duas enfermeiras auditoras (validação dos itens cobrados e definição dos protocolos baseados em regras) e um analista da Central de Pagamentos — CP (exportação dos dados do sistema de business intelligence).

Processo prévio à automação

Antes da introdução do sistema automatizado, a auditoria das contas era realizada manualmente por uma enfermeira auditora. Os dados eram extraídos de um banco corporativo no Microsoft Power BI, exportados em planilhas e avaliados linha a linha. Esse processo demandava aproximadamente dois dias úteis para analisar cerca de 2.000 autorizações semanais, apresentava risco de erros operacionais e dificultava a padronização dos critérios adotados. A ausência de mecanismos automatizados também limitava a rastreabilidade das decisões.

Fonte de dados e pré-processamento

Foram utilizados dados de pagamentos realizados a prestadores hospitalares, exportados do sistema de BI já desprovidos de identificadores diretos (nome, CPF ou dados de contato). Os campos disponíveis incluíam: número da autorização (handle), código e descrição dos procedimentos segundo a Terminologia Unificada da Saúde Suplementar (TUSS), tabela de origem, quantidade autorizada e executada, valor pago, datas de realização e de alta, CID principal, sexo e data de nascimento. Do total bruto de aproximadamente 2.010 linhas, foram removidos registros duplicados (~5%), resultando em 1.910 linhas correspondentes a 1.909 autorizações únicas (handles). Considerou-se duplicidade a coincidência de número de autorização (handle) e código TUSS no mesmo evento, mantendo-se um único registro por par handle–TUSS. A unidade de análise foi a autorização (handle); como um mesmo prestador e um mesmo paciente podem responder por múltiplas autorizações, há dependência entre observações, o que recomenda cautela na interpretação dos intervalos de confiança e foi considerado uma limitação. A janela analisada (quatro dias, de 26 a 29 de maio de 2025) corresponde a uma semana típica de faturamento hospitalar encaminhado à regulação; por se tratar de período curto, não capta variações sazonais nem mudanças de prestadores ou de regras contratuais, conforme discutido nas limitações.

Estrutura do Robô-Glosa

A versão 5.7 da ferramenta foi desenvolvida em linguagem VBA integrada ao Microsoft Excel, com arquitetura modular e parametrizável. A versão incorpora melhorias acumuladas de ciclos anteriores de desenvolvimento, incluindo a adição de regras contratuais específicas por prestador e a detecção estatística de valores anômalos. O pipeline de processamento compreende oito etapas: (1) ingestão e mapeamento dinâmico de colunas; (2) carregamento de parâmetros de referência (pacotes, OPME, drogas de alto custo, limites) a partir de aba dedicada; (3) transferência dos dados para array bidimensional em memória RAM; (4) pré-passada para marcação de internações com pacote e geração de hash por autorização e TUSS; (5) aplicação do motor de 11 regras de conformidade; (6) cálculo do valor estimado de glosa por linha; (7) escrita dos resultados com sinalização visual; e (8) geração de resumo quantitativo. O código-fonte não está disponível publicamente por conter regras contratuais de propriedade da operadora, porém, a estrutura lógica e os parâmetros são descritos neste artigo de forma a permitir a replicação em outros contextos.

Regras de conformidade implementadas

O sistema implementa 11 regras de conformidade, organizadas em três domínios. No domínio contratual: (1) alerta de cobrança indevida em pacotes, quando procedimentos são lançados isoladamente em eventos com tabela de pacote; (2) regras contratuais específicas por prestador, parametrizáveis por meio de listas auxiliares mantidas pela equipe de auditoria. No domínio técnico-administrativo: (3) duplicidade de procedimento (mesmo handle + TUSS); (4) quantidade paga superior à autorizada; (5) procedimento realizado após data de alta; (6) inconsistência entre honorários e procedimento; (7) implantes ou OPME duplicados; (8) incompatibilidade entre procedimento de alto custo e CID; (9) incompatibilidade idade/sexo; (10) cobrança de taxa de sala em procedimentos de porte 0 ou 1; (11) fracionamento excessivo de materiais ou medicamentos. Os códigos de procedimentos utilizados pelo sistema seguem a Terminologia Unificada da Saúde Suplementar, conforme estabelecido no Rol de Procedimentos e Eventos em Saúde (BRASIL, 2021). No domínio estatístico, o sistema inclui a detecção de valores unitários (outliers), aplicando uma lógica incremental (média ± 2 desvios-padrão por código TUSS), calculada em uma única passada, sem necessidade de segunda varredura dos dados. A especificação operacional de cada uma das 11 regras (condição lógica, fonte de parametrização, limiar e exemplos) é apresentada de forma anonimizada no Quadro Suplementar 1, sem qualquer elemento que permita identificar prestadores ou a operadora. Quanto à detecção de valores anômalos, o critério média ± 2 desvios-padrão por código TUSS foi adotado por simplicidade computacional; reconhece-se, contudo, que os valores unitários hospitalares raramente seguem distribuição normal, de modo que medidas mais robustas a distribuições assimétricas (mediana $\pm$ intervalo interquartilico ou percentis 95/99) tendem a ser mais adequadas e estão previstas para os próximos ciclos. A regra estatística foi aplicada apenas a códigos TUSS com número de lançamentos suficiente para estimativa estável da média e do desvio-padrão no período, descartando-se códigos com poucos registros.

Avaliação de desempenho

A unidade de análise foi a autorização (handle), refletindo a lógica operacional efetiva da auditoria. O desempenho do robô foi avaliado por meio da comparação de suas classificações com as da auditoria manual realizada pela enfermeira auditora, adotada como referência pragmática. Verdadeiro positivo (VP) foi definido como autorização glosada por ambos os sistemas; falso negativo (FN) como autorização glosada pela enfermeira, mas não pelo robô; falso positivo (FP) como autorização glosada pelo robô, mas não pela enfermeira; e verdadeiro negativo (VN) como autorização aprovada por

ambos. Foram calculados a acurácia $[(VP+VN)/N]$, a sensibilidade $[VP/(VP+FN)]$, a especificidade $[VN/(VN+FP)]$, o valor preditivo positivo (VPP) $[VP/(VP+FP)]$ e o valor preditivo negativo (VPN) $[VN/(VN+FN)]$. A métrica anteriormente denominada “precisão” foi renomeada para valor preditivo positivo (VPP): no sentido metrológico-epidemiológico, precisão designa reprodutibilidade (concordância intra e interobservador), enquanto $VP/(VP+FP)$ corresponde ao VPP. Como o conjunto analisado é pré-selecionado, com alta prevalência de glosa, o VPP e o VPN dependem da prevalência e não devem ser generalizados para o universo de contas da operadora. Os intervalos de confiança de 95% foram estimados pelo método de Wilson para todas as proporções, por seu melhor desempenho em proporções extremas em comparação ao método de Wald; para as métricas iguais a 100% (especificidade e VPP), reportou-se adicionalmente o intervalo exato de Clopper-Pearson. Os cálculos foram realizados em Python 3.11, com a função `proportion_confint` do pacote `statsmodels`. Cabe registrar que a auditoria manual foi conduzida por avaliadora única e não cega e que o robô aplica um subconjunto das mesmas regras empregadas por essa auditora, o que configura circularidade parcial do desenho, detalhada a seguir e retomada nas limitações.

Padrão de referência e controle de viés

A auditoria manual da enfermeira auditora foi adotada como padrão de referência pragmático. A classificação manual baseou-se nos mesmos critérios clínicos, contratuais e administrativos posteriormente parametrizados no robô, aplicados linha a linha sobre a planilha exportada. A revisão foi realizada por uma única auditora, sem dupla revisão independente e sem cegamento formal em relação ao resultado do robô; a auditoria manual e a execução automatizada incidiram sobre o mesmo conjunto de dados, com intervalo de poucos dias entre as etapas. As discordâncias entre o robô e a auditora foram classificadas como falsos negativos ou falsos positivos, conforme as definições adotadas, sem reconciliação por terceiro revisor. O desenvolvedor principal (médico auditor) participou tanto da programação quanto da análise dos resultados, o que constitui potencial viés; as duas enfermeiras auditoras atuaram na validação dos itens e na definição dos protocolos, e o analista da Central de Pagamentos, na exportação dos dados. Reconhece-se a limitação de avaliador único e não cego, recomendando-se, em validações futuras, a adoção de padrão de referência composto, com dupla revisão independente por auditores que não tenham participado do desenvolvimento.

Aspectos éticos

Este projeto caracteriza-se como avaliação administrativa interna, utilizando exclusivamente dados exportados sem identificadores diretos, sem possibilidade de identificação individual de pacientes ou profissionais, dispensando a submissão ao Comitê de Ética em Pesquisa. Todo o processo foi conduzido em conformidade com a Lei Geral de Proteção de Dados (BRASIL, 2018a). A dispensa de apreciação por Comitê de Ética em Pesquisa fundamenta-se na Resolução CNS nº 510/2016, que dispensa registro e avaliação por CEP de pesquisas que utilizam dados não identificados, bem como nos princípios da Resolução CNS nº 466/2012, no que couber. A operadora concedeu autorização institucional formal para a análise e para a divulgação dos resultados de forma agregada e anonimizada. Não são apresentados elementos — denominação, localização específica, prestadores, valores contratuais ou parâmetros — que, isolados ou combinados, permitam identificar a operadora ou seus prestadores.

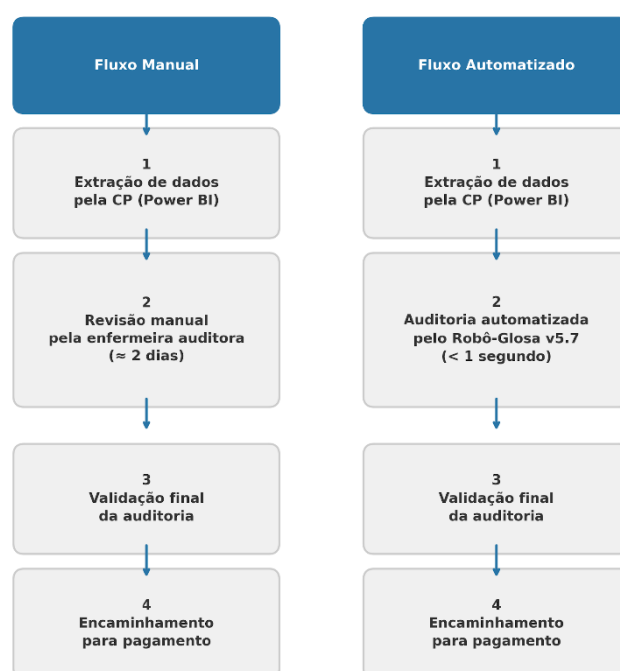


Figura 1. Fluxograma comparativo dos fluxos de trabalho de auditoria manual e automatizada. A revisão manual pela enfermeira auditora consome aproximadamente 2 dias, enquanto o processo automatizado pelo Robô-Glosa v5.7 é concluído em menos de 1 segundo.

RESULTADOS

No período de 26 a 29 de maio de 2025, o Robô-Glosa processou 1.910 linhas de faturamento hospitalar, correspondentes a 1.909 autorizações únicas (handles), em menos de 1 segundo de processamento computacional, excluídas as etapas prévias de exportação

e de pré-processamento dos dados. Em contraste, o mesmo volume de dados exigia aproximadamente dois dias úteis para a revisão manual.

Na comparação com a auditoria manual, o sistema identificou corretamente 1.430 autorizações como passíveis de glosa (verdadeiros positivos) e 378 como aprovadas (verdadeiros negativos), sem produzir nenhum falso positivo (Figura 2). Foram identificados 101 falsos negativos, correspondentes a autorizações glosadas pela enfermeira, mas não detectadas pelo robô. As métricas de desempenho, com intervalos de confiança de 95% pelo método de Wilson, foram: acurácia de 94,71% (IC 95%: 93,61–95,63), sensibilidade de 93,40% (IC 95%: 92,05–94,54), especificidade de 100% (IC 95%: 98,99–100; Clopper-Pearson 99,03–100), valor preditivo positivo (VPP) de 100% (IC 95%: 99,73–100; Clopper-Pearson 99,74–100) e valor preditivo negativo (VPN) de 78,91% (IC 95%: 75,04–82,33). O VPP e o VPN refletem a alta prevalência de glosa do conjunto pré-selecionado (80,2%) e não devem ser extrapolados para o universo de contas da operadora.

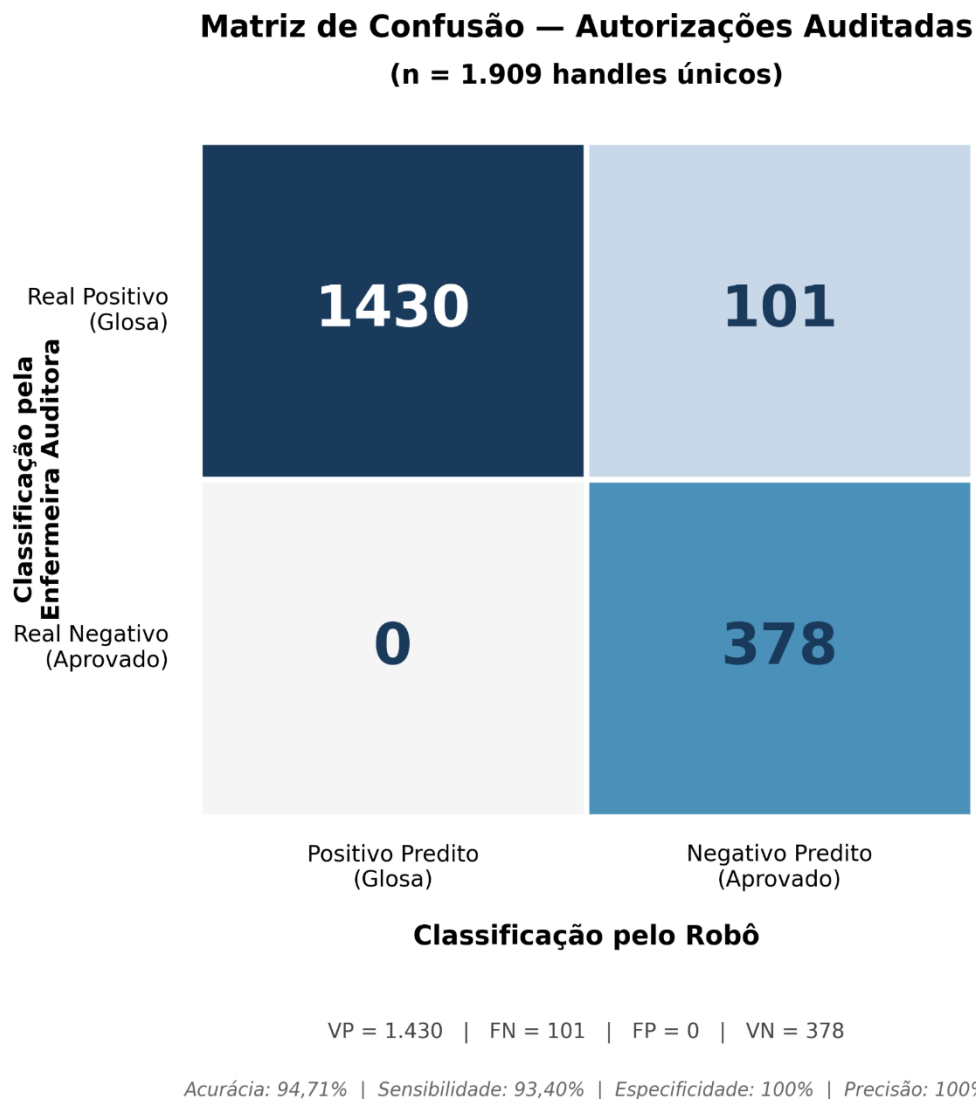


Figura 2. Matriz de confusão do Robô-Glosa comparada à auditoria manual da enfermeira (n = 1.909 handles únicos, no período de 26 a 29 de maio de 2025). VP = verdadeiro positivo; FN = falso negativo; FP = falso positivo; VN = verdadeiro negativo.

A prevalência de glosa no conjunto analisado foi de 80,2% (1.531/1.909 handles). Esse valor reflete a pré-seleção do fluxo analisado — contas hospitalares de maior complexidade encaminhadas ao setor de regulação para revisão detalhada — e não deve ser interpretado como a prevalência de inconformidades no universo total de contas da operadora. A principal categoria de glosa identificada pelo robô foi a cobrança indevida de itens já incluídos em pacotes assistenciais, seguida da cobrança de taxa de sala em procedimentos de porte 0 ou 1. A distribuição quantitativa das glosas por regra não foi registrada individualmente nesta validação preliminar, o que constitui um dos aprimoramentos previstos para os próximos ciclos de avaliação.

A equipe de auditoria avaliou que os 101 falsos negativos (6,6% das glosas reais) se concentraram em situações que exigiam julgamento clínico contextual ou interpretação de nuances contratuais não parametrizáveis por regras estáticas, como cláusulas de exceção por prestador ou avaliação da pertinência clínica de materiais especiais. Essa análise indicou os limites inerentes a um sistema baseado exclusivamente em regras e apontou as áreas que mais se beneficiariam de complementação por modelos de aprendizado de máquina ou por revisão manual voltada aos casos não cobertos pelas regras parametrizáveis.

DISCUSSÃO

Os resultados desta validação preliminar indicam que o Robô-Glosa, uma ferramenta de automação baseada em regras desenvolvida inteiramente em VBA/Excel, alcançou desempenho comparável ao da revisão manual especializada na identificação de inconformidades em contas hospitalares, com acurácia de 94,71% e sensibilidade de 93,40%. Dois achados merecem destaque: a ausência absoluta de falsos positivos e o perfil dos 101 falsos negativos. Ambos revelam não apenas o desempenho do sistema, mas também as características intrínsecas de uma abordagem baseada em regras quando aplicada à auditoria de contas médicas.

O valor preditivo positivo (VPP) e a especificidade de 100%, nenhuma autorização legítima indevidamente sinalizada para glosa, constituem, à primeira vista, um resultado excepcional. Entretanto, é necessário interpretar esse achado à luz do desenho do estudo. O Robô-Glosa opera exclusivamente de acordo com regras predefinidas pela própria equipe de auditoria, que é a mesma utilizada como referência na validação. Por construção, o sistema aplica um subconjunto da lógica empregada pela enfermeira auditora na revisão manual: não introduz critérios novos nem realiza julgamentos que extrapolem os parâmetros codificados. Nesse sentido, a ausência de falsos positivos é, em grande medida, consequência do desenho, e não uma propriedade diagnóstica independente. Essa distinção é importante porque a garantia de que o robô não gera glosas injustificadas tem relevância prática direta: preserva a relação de confiança entre a operadora e os prestadores, aspecto reconhecido como crítico para a sustentabilidade contratual na saúde suplementar (GAMARRA, 2018). Contudo, impede que se interprete o VPP de 100% como evidência de superioridade diagnóstica do sistema.

Os 101 falsos negativos (6,6% das glosas reais) delineiam com clareza os limites de um sistema determinístico. Conforme avaliado pela equipe de auditoria, esses casos

concentraram-se em situações dependentes de julgamento clínico contextual — exceções contratuais não codificadas, avaliação de pertinência de materiais especiais — que, por natureza, não são capturáveis por regras estáticas. Essa constatação é coerente com a literatura. Hanskamp-Sebregts et al. (2019), ao avaliarem sistemas de auditoria de segurança do paciente, concluíram que a efetividade dessas ferramentas depende diretamente de quanto elas espelham as necessidades e os protocolos específicos de cada contexto institucional. As 11 regras implementadas no Robô-Glosa cobrem os cenários mais frequentes e padronizáveis, mas não substituem — nem pretendem substituir — o julgamento clínico do auditor especializado. Nesse sentido, a ferramenta posiciona-se como instrumento de pré-triagem: ao sinalizar automaticamente as inconformidades parametrizáveis, libera o auditor humano para concentrar-se nos casos que exigem interpretação contextual, elevando a qualidade global do processo.

A comparação com experiências anteriores de automação na auditoria de contas no Brasil reforça a contribuição deste relato. Baurakiades et al. (2012) descreveram a implementação de um sistema automatizado em hospital privado, demonstrando ganhos de produtividade; contudo, não apresentaram métricas formais de desempenho diagnóstico (acurácia, sensibilidade ou especificidade), o que limita a comparabilidade direta e evidencia a escassez de validações quantitativas na literatura brasileira sobre o tema. Araya, Pinheiro e Bezerra (2016) propuseram um sistema baseado em mineração de dados para auditoria hospitalar, alcançando resultados promissores na detecção de inconsistências, mas dependente de uma plataforma computacional dedicada e de expertise em programação. No cenário internacional, Chen et al. (2024), ao aplicarem Robotic Process Automation (RPA) combinada com Lean Six Sigma na auditoria de contas hospitalares em Taiwan, reduziram o tempo de processamento em 380 minutos e elevaram a eficiência do ciclo de 69% para 95%, validando o potencial da automação baseada em regras em contextos de alta complexidade regulatória. O Robô-Glosa se aproxima conceitualmente dessa abordagem — automação “de fora para dentro”, sem alterar os sistemas de informação originais —, mas distingue-se pela acessibilidade radical: opera no Microsoft Excel, dispensa infraestrutura adicional e permite que a própria equipe de auditoria parametrize e atualize as regras por meio de listas auxiliares em planilha, sem necessidade de alteração de código.

Essa acessibilidade posiciona o Robô-Glosa em um nicho específico no espectro de soluções disponíveis para a auditoria de contas médicas. De um lado, a revisão manual pura, ainda predominante em operadoras de menor porte, é lenta e sujeita à variabilidade

subjetiva entre auditores (CRISTINA et al., 2021; SCARPARO et al., 2009). Do outro lado, plataformas avançadas de inteligência artificial, que incorporam aprendizado de máquina, processamento de linguagem natural e integração com prontuários eletrônicos, oferecem capacidade analítica superior, mas demandam investimento substancial em dados estruturados, infraestrutura computacional e equipe técnica especializada (NIMKAR et al., 2024). O Robô-Glosa ocupa o espaço intermediário: automação determinística de baixo custo, com alta precisão em cenários parametrizáveis e limitação explícita em cenários que requerem julgamento. Para operadoras de autogestão, que dispõem de estruturas enxutas e de recursos tecnológicos limitados (UNIDAS, 2025), essa solução intermediária pode representar o ponto de entrada mais pragmático para a automação da auditoria.

Para além da redução no tempo de processamento — de aproximadamente dois dias para menos de 1 segundo —, a automação trouxe ganhos operacionais qualitativos que merecem destaque. A padronização dos critérios de glosa elimina a variabilidade subjetiva entre auditores, um problema frequentemente apontado na literatura (SCARPARO et al., 2009). A rastreabilidade das decisões é ampliada, uma vez que cada sinalização está vinculada à regra específica que a originou, mitigando uma fragilidade reconhecida nos processos manuais de auditoria hospitalar (RODRIGUES; PERROCA; JERICÓ, 2004). A escalabilidade do sistema também merece menção: o mesmo robô pode ser aplicado a múltiplos prestadores sem custo marginal significativo, bastando a atualização dos parâmetros contratuais. Esses atributos convergem com as diretrizes da ANS para o fortalecimento da governança nas operadoras de autogestão (BRASIL, 2025b) e, simultaneamente, alinham-se à Resolução CFM nº 2.448/2025, que reafirmou a auditoria médica como ato privativo do médico e vedou a substituição do juízo clínico por critérios exclusivamente administrativos ou financeiros (BRASIL, 2025c). O Robô-Glosa, ao operar como instrumento de pré-triagem e não como decisor final, respeita essa delimitação: sinaliza potenciais inconformidades com base em regras técnicas e contratuais, mas reserva ao auditor médico a palavra final sobre a pertinência clínica da glosa.

A comparação com o processo manual exige, contudo, ressalvas metodológicas. A auditoria manual pela enfermeira, adotada como referência pragmática, apresenta limitações próprias reconhecidas na literatura, incluindo variabilidade intra- e inter-avaliadora e vulnerabilidade à fadiga cognitiva em tarefas repetitivas (CRISTINA et al., 2021). Portanto, parte dos falsos negativos identificados pode corresponder a glosas

discutíveis no processo manual, e não necessariamente a inconformidades reais não capturadas pelo robô. Adicionalmente, o desenvolvedor principal do sistema participou da análise dos resultados, o que configura potencial viés na avaliação. Estudos futuros de validação formal devem adotar um padrão de referência composto, com dupla revisão independente por auditores que não tenham participado do desenvolvimento, de modo a avaliar de forma mais robusta tanto a sensibilidade do sistema quanto a confiabilidade da própria auditoria manual.

Limitações adicionais devem ser consideradas. A validação foi conduzida com dados de um único período de quatro dias de faturamento hospitalar, o que impediu a avaliação do desempenho diante de variações sazonais, mudanças de prestadores ou atualizações contratuais. A análise de impacto financeiro não foi incluída por envolver valores contratuais confidenciais entre a operadora e os prestadores credenciados. Embora a identificação de inconformidades tenha implicação direta na contenção de custos, os autores reconhecem esta limitação e consideram que estudos futuros, com autorização institucional formal, devem quantificar o impacto econômico da automação na gestão financeira da operadora. O sistema depende da exportação manual de dados do Power BI e não se integra a prontuários eletrônicos ou sistemas de faturamento em tempo real, restringindo a automação plena do processo — limitação compartilhada pela maioria das soluções de RPA em saúde atualmente disponíveis (CHEN et al., 2024). A distribuição quantitativa das glosas por regra não foi registrada nesta validação preliminar, o que limita a análise da contribuição relativa de cada critério de conformidade.

Apesar dessas limitações, o projeto oferece potencial de replicabilidade relevante para outras operadoras de autogestão e para unidades de auditoria hospitalar. Como próximos passos, planeja-se a expansão da base de validação para múltiplas semanas e prestadores, o registro individualizado das glosas por regra, a integração com dashboards de BI em tempo real e, em horizonte mais amplo, a incorporação de modelos de aprendizado supervisionado voltados à redução dos falsos negativos em cenários que exigem julgamento clínico contextual (COELHO, 2025).

CONCLUSÃO

O Robô-Glosa demonstrou viabilidade como ferramenta de pré-triagem automatizada de contas médicas hospitalares em uma operadora brasileira de autogestão. O sistema combinou alto valor preditivo positivo (VPP) — sem nenhum falso positivo —

com sensibilidade de 93,40% na detecção de inconformidades, reduzindo o tempo de análise de dias para segundos e eliminando a variabilidade subjetiva entre auditores.

Três contribuições deste relato merecem destaque. Em primeiro lugar, a validação com métricas formais de desempenho diagnóstico e intervalos de confiança preenche uma lacuna na literatura brasileira sobre automação na auditoria de contas, na qual experiências anteriores relataram ganhos qualitativos, sem quantificação padronizada. Em segundo lugar, a análise dos 101 falsos negativos não apenas delimitou os cenários em que regras estáticas são insuficientes, mas também indicou com precisão as áreas que mais se beneficiariam de complementação por modelos de aprendizado de máquina ou por revisão manual dirigida. Em terceiro lugar, a demonstração de que uma ferramenta operando inteiramente em VBA/Excel pode alcançar esse patamar de desempenho reforça que a barreira para a automação da auditoria em operadoras de menor porte não é tecnológica, mas organizacional — reside na padronização dos critérios, na colaboração entre equipes clínicas e administrativas e na disposição institucional para incorporar processos baseados em evidência.

Ao operar como instrumento de pré-triagem e não como decisor final, o Robô-Glosa preserva a centralidade do julgamento clínico do auditor médico, alinhando-se às diretrizes recentes do Conselho Federal de Medicina sobre os limites da automação na auditoria em saúde. Essa delimitação, longe de ser uma fraqueza, define o espaço legítimo e sustentável de atuação de ferramentas baseadas em regras no ecossistema da saúde suplementar brasileira.

Conflito de interesses

O autor declara não haver conflito de interesses relacionado a este trabalho.

Financiamento

Este estudo não recebeu financiamento externo.

Referências

1. ARAYA, A.; PINHEIRO, A. L. S.; BEZERRA, L. C. A. An intelligent system for medical auditing in hospital bills. *Journal of Health Informatics*, v. 8, n. 2, p. 49-56, 2016.
2. BAURAKIADES, E. et al. Automated medical billing auditing system: a case study in a private hospital. *Journal of Health Informatics*, v. 4, n. 1, p. 15-21, 2012.
3. BRASIL. Lei nº 13.709, de 14 de agosto de 2018. Dispõe sobre a proteção de dados pessoais (LGPD). *Diário Oficial da União*, Brasília, DF, 15 ago. 2018a.
4. BRASIL. Agência Nacional de Saúde Suplementar. Resolução Normativa nº 465, de 24 de fevereiro de 2021. Dispõe sobre o Rol de Procedimentos e Eventos em Saúde no âmbito da Saúde Suplementar. *Diário Oficial da União*, Brasília, DF, 25 fev. 2021.
5. BRASIL. Agência Nacional de Saúde Suplementar. Resolução Normativa nº 623, de 2024. Dispõe sobre as regras do relacionamento entre operadoras e beneficiários de planos de saúde. Rio de Janeiro: ANS, 2024.
6. BRASIL. Agência Nacional de Saúde Suplementar. Dados Gerais do Setor de Saúde Suplementar. Rio de Janeiro: ANS, 2025a.
7. BRASIL. Agência Nacional de Saúde Suplementar. Resolução Normativa nº 649, de 31 de outubro de 2025. Dispõe sobre as regras aplicáveis às operadoras de autogestão. Rio de Janeiro: ANS, 2025b.
8. BRASIL. Conselho Federal de Medicina. Resolução nº 2.448, de 23 de outubro de 2025. Dispõe sobre normas específicas para a auditoria médica e revoga a Resolução CFM nº 1.614/2001. *Diário Oficial da União*, Brasília, DF, 4 nov. 2025c.
9. CHEN, Y. H. et al. A case study of lean digital transformation through robotic process automation in healthcare. *Scientific Reports*, v. 14, n. 14617, 2024.
10. COELHO, G. Auditoria baseada em dados: classificação de risco estatístico em prestadores de saúde com base no faturamento e volume de atendimentos. *Brazilian Journal of Business*, v. 7, n. 2, e79268, 2025.
11. CRISTINA, F. et al. Aplicação da mineração de dados na auditoria de contas médicas hospitalares. *Journal of Health Informatics*, v. 13, n. 2, p. 31-36, 2021.
12. GAMARRA, T. P. N. Auditoria na saúde suplementar: uma revisão integrativa. *Revista de Gestão em Sistemas de Saúde*, v. 7, n. 3, p. 221-237, 2018.

13. HANSKAMP-SEBREGTS, M. et al. Effects of patient safety auditing in hospital care: results of a mixed-method evaluation. *International Journal for Quality in Health Care*, v. 31, n. 7, p. G8-G15, 2019.
14. NIMKAR, A. V. et al. Increasing trends of artificial intelligence with robotic process automation in health care: a narrative review. *Cureus*, v. 16, n. 9, e69680, 2024.
15. PINHEIRO, A. L. S. et al. Health management: the use of information systems and knowledge sharing for the decision making process. *Texto & Contexto – Enfermagem*, v. 25, n. 3, e3440015, 2016.
16. RODRIGUES, V. A.; PERROCA, M. G.; JERICÓ, M. C. Glosas hospitalares: importância das anotações de enfermagem. *Arquivos de Ciências da Saúde*, v. 11, n. 4, p. 165-171, 2004.
17. RULE, A.; CHIANG, M. F.; HRIBAR, M. R. Using electronic health record audit logs to study clinical activity: a systematic review of aims, measures, and methods. *Journal of the American Medical Informatics Association*, v. 27, n. 3, p. 480-490, 2020.
18. SCARPARO, A. F. et al. Abordagem conceitual de métodos e finalidade da auditoria de enfermagem. *Revista da Rede de Enfermagem do Nordeste*, v. 10, n. 1, p. 124-130, 2009.
19. UNIDAS. União Nacional das Instituições de Autogestão em Saúde. *Panorama da Autogestão em Saúde*. São Paulo: UNIDAS, 2025.