

Using AI for Developing Performance Audit Reports

Artificial Intelligence (AI) is uniquely positioned to help expedite our audit report writing process. This document is intended to serve as a guide for auditors wanting to incorporate AI into the writing process.

Contents

Converting a PFR Document into a Report Outline	1
Developing Sections of the Report Using the Outline.....	2
Condensing Report Sections.....	3
Modify Report Sections for Tone	5

Converting a PFR Document into a Report Outline

The concept here is simple: we upload our PDF document into ChatGTP and ask it to create an outline of the content for us.

Step 1: Load a PFR document into ChatGTP. Simply drag and drop a copy of your PFR document into ChatGTP. Once loaded, you can simply ask ChatGTP to create an outline for you.

Step 2: Instruct ChatGTP to create an outline for you. This can be done in numerous ways, and it's very easy to modify your prompt to get results more aligned with what you're looking for. Here's a few examples of the types of results ChatGTP generates based on a PFR document (I used the PFR for our Medicaid Lead Testing audit, published in December 2023):

Prompt: Develop an outline of this document in 750 words or less at a fourteenth grade reading level including these sections: background, audit findings, recommendations, and conclusion.

Results: Here's a short snippet of what ChatGTP generated based on my prompt:

Commented [JS1]: Send Michael Eglinski a copy of this when finished.

Commented [GS2]: This is clear and helpful guidance. The examples are great too. I had only one thought/suggestion similar to one of Kayvon's so I tacked it onto his.

Commented [FT4]: Agreeing with all the comments about how useful this will be.

This may be outside the realm of what you asked, but any thoughts on having AI draft report recommendations?

Commented [JS5]: All of my experimentation was done using ChatGTP. If we end up using CoPilot (very likely), we'll need to update this guidance document accordingly.

Commented [FS6R5]: It'll be Copilot. Maybe ChatGPT down the road, but not soon.

Also, slight correction. GPT not GTP.

Commented [FT7]: This comment might be better in a different context, but a question that comes to mind for me is this: What do you do with PFRs that aren't fully baked?

I think this step you have makes complete sense, but I'm trying to align it with the mechanics of our existing process.

On the day to day, I imagine there would need to be some sort of filtering process.

Sorry this isn't super coherent, and I might not be really clearly understanding the purpose of the guide anyway...

Background:

- **Lead Exposure and Health Impact:**

- Lead exposure poses significant risks to children, particularly affecting brain and nervous system development.
- The CDC uses a blood lead reference value of 3.5 µg/dL to identify children with elevated lead levels, necessitating intervention.

- **Medicaid and Lead Testing:**

- Federal law mandates lead screening for children enrolled in Medicaid at 12 and 24 months under the Early and Periodic Screening, Diagnostic, and Treatment (EPSDT) benefit.
- Despite this requirement, testing rates remain low, leading to potential undiagnosed and untreated cases of lead poisoning.

Audit Findings:

1. **Low Lead Testing Rates:**

- Only 13% of Medicaid-enrolled children in Washington receive the mandated lead tests at both 12 and 24 months.
- Factors contributing to low testing rates include lack of awareness among parents and providers, reluctance from families, and insufficient reimbursement rates for lead testing.
- The low testing rate is alarming because it means many children with elevated blood lead levels may not be identified and treated promptly.

2. **Inadequate Monitoring and Oversight:**

- Washington's current system for monitoring lead testing compliance is deficient.
- There are no specific performance measures or systematic analyses to track and improve lead testing rates among Medicaid children.
- This lack of oversight can result in continued non-compliance and missed opportunities for early intervention.

Note: I modified facts/figures in the PFR to sanitize it; the results won't necessarily tie to our actual report.

Commented [KZ8]: Might be worth indicating whether sanitization is an expectation. I'm guessing the answer is no with CoPilot, but I could see someone assuming they need to do so based on this

Step 3: Modify your prompt to enhance your results. AI tools are easily able to modify the results produced based on very simple tweaks to your prompt. Some potential tweaks you can consider include changes to:

- The maximum word total allowed.
- The desired grade reading level.
- The tone (neutral, evaluative, instructional, etc.)
- Required sections (if any).

Developing Sections of the Report Using the Outline

Once we have an outline generated by ChatGTP, we can ask it to more fully develop content based on that outline, still using the raw information from the PFR. Here's how:

Step 1: In a new chat thread, upload the PFR document to ChatGTP.

Commented [FS9]: Any reason why you're not using the same thread? It will already have the PFR document. (Maybe you tested this and it didn't work.)

Step 2: Instruct ChatGTP to use the PFR to develop content based on the outline.

Prompt: This document summarizes our preliminary findings and recommendations for an audit of lead testing for Medicaid-eligible children. Here is an outline of some key points I'd like to develop:

Low Lead Testing Rates

- Only 13% of Medicaid-enrolled children in Washington receive the mandated lead tests at both 12 and 24 months.
- Factors contributing to low testing rates include lack of awareness among parents and providers, reluctance from families, and insufficient reimbursement rates for lead testing.
- The low testing rate is alarming because it means many children with elevated blood lead levels may not be identified and treated promptly.

Here are my instructions

1. Please flesh out the three bullet points into individual paragraphs, drawing on relevant information from the preliminary findings document. In preparing these paragraphs, make sure that each paragraph is written in a deductive manner, each paragraph has a clear topic sentence, and all other sentences in a given paragraph directly relate to, or support, the topic sentence.
2. Critique your results from step 1. Identify ways to improve the writing so they paragraphs are clear, concise, and convincing.
3. Revise your paragraphs from step 1 to address the critique from step 2. Please be sure to show all of your work, so I can see the results of each step.

Results (I manually bolded the first sentence in each paragraph):

Washington is not meeting federal Medicaid requirements for blood lead testing, with only 13% of enrolled children receiving tests at both 12 and 24 months. This compliance gap leaves a large number of children unscreened during critical early development stages. Data shows that testing rates have remained low across multiple cohorts, signaling a persistent and systemic problem. Without intervention, thousands of children may miss the opportunity for early detection and treatment of lead exposure.

Several barriers contribute to Washington's low lead testing rates among Medicaid-enrolled children. Many providers are unaware of the testing requirement or do not view lead exposure as a pressing concern in their patient population. Others forgo testing when parents resist, or when more immediate health issues take priority during visits. Structural issues compound these challenges—low reimbursement rates, difficulty scheduling well-child visits, and lack of clarity around Medicaid eligibility all contribute to inconsistent screening. Together, these factors create significant gaps in testing coverage.

Failing to test Medicaid-enrolled children for lead exposure has serious and lasting consequences. Without screening, children with elevated blood lead levels may go undetected and untreated, putting them at risk for developmental delays, cognitive impairment, and behavioral issues. Lead exposure can silently impair a child's ability to succeed in school and in life. Because symptoms are rarely visible early on, blood testing is the only reliable way to identify those in need of help. Increasing testing rates is essential to protect children's health and potential.

Commented [KZ10]: Does this need to be done one section at a time to work well, or could the whole outline and PFR be applied at the same time? If the latter was workable I would think it would be more efficient, but not sure if this is a technical limitation.

Commented [SM11]: It's interesting that we have to ask it to do this as a second step. Couldn't we have included this in the step one instructions, for the paragraphs to be clear, concise and convincing?

Commented [FS12R11]: Either way probably works, but I think Melissa's suggestion is probably more direct.

Commented [KZ13]: It might be interesting to actually see a snippet of what this looks like to help folks understand a little but more about what's going on under the hood.

Condensing Report Sections

Following the Director's Review of the draft, teams are often asked to condense certain sections of the report (i.e. please cut this content by about 50% for the final report). Here's how ChatGTP can help with that process.

Commented [SM14]: This is awesome! It will be great to use ChatGTP to do this.

Step 1: Ask Chat GTP to condense a section of your report by some amount.

Prompt: I'm going to include a current section of an audit report I'm writing that I need you to condense. Here's my instructions:

1. Please condense the current section by 50%. When condensing the section, please use a twelfth-grade reading level.
2. Critique your results from step 1. Identify ways to improve the writing so they paragraphs are clear, concise, and convincing.
3. Revise your paragraphs from step 1 to address the critique from step 2.
Please be sure to show all of your work, so I can see the results of each step.

Current Section

Additionally, HCA has not used its existing performance measure to actively monitor or increase testing compliance

In 2021, MCOs tested between 27 percent and 41 percent of the children in their care at least once, as measured by HCA's existing performance measure. For a test that is to be universally applied, these are low rates. HCA has left it up to MCOs to determine their next steps for addressing the lead testing performance measure, and has not required any MCO to take corrective action for underperforming.

Other states have taken more direct approaches to improving provider compliance. For example, Wisconsin, New York and North Carolina all use billing data to identify low-performing providers and then notify them on how well they are or are not meeting the Medicaid requirement. HCA could require the MCOs to do something similar with their provider groups. Wisconsin also used Medicaid billing data and blood lead testing data to identify Medicaid-enrolled children who had not received the required tests. Providers were notified of children in their practice who had not been tested. After implementing this provider outreach, the share of children who received an age-appropriate test increased from 46 percent to 55 percent. As HCA works to improve its systems of monitoring and accountability, the agency could consider how best to work with the MCOs to give their providers more direct feedback on their own performance.

Results (I manually bolded the first sentence):

HCA has not actively used its lead testing performance measure to improve compliance. In 2021, only 27% to 41% of children covered by MCOs received at least one test—well below expectations for a universally required screening. Despite this, HCA has not required MCOs to correct poor performance. Other states have taken stronger steps. Wisconsin, New York, and North Carolina use billing data to flag providers with low testing rates and notify them of their performance. Wisconsin also identified Medicaid-enrolled children who missed required tests and informed their providers directly. This outreach raised testing rates from 46% to 55%. HCA could adopt similar strategies by working with MCOs to give providers clear, timely feedback and hold them accountable for meeting testing goals.

Commented [KZ15]: Wondering if you've found any additional prompts helpful to provide parameters for this condensing (i.e. condense by simplifying sentence structure, choosing plainer language, etc.) or whether it's been better to just let it decide on its own (as you did here)?

Commented [GS16R15]: I had a similar wondering. With this prompt, I could see a situation where ChatGPT cut some content that was important so I was wondering if there is an additional prompt that would help reduce that risk.

Modify Report Sections for Tone

It's not uncommon for us to occasionally address sensitive topics or issues in our reports. Here's how ChatGTP can be used to modify report sections for tone.

Step 1: Ask ChatGTP to modify the tone of your language to make it less abrasive.

Prompt: I'm going to give you a paragraph we're working on for a government performance audit of the Department of Food Inspections. The audit team developed this paragraph to communicate one of its audit findings in an audit report we're developing. However, as a manager in our audit office, I'm concerned that the tone of the paragraph is a little too abrasive and accusatory.

I'd like you to revise this paragraph to convey the same general message, but in a more neutral, balanced, and objective tone. In making your revision, please make sure you develop a strong thesis statement that is supported or explained by the subsequent sentences in the paragraph.

Original Paragraph

The Department of Food Inspections has failed to conduct timely restaurant inspections, which puts members of the public at significant risk for food-borne illnesses. We found the Department only completed about 70% of its restaurant inspections on time during fiscal year 2025. Although the Department told us that was primarily because it was still catching up from a backlog of inspections created during the COVID-19 pandemic. We asked the Department whether they had requested additional funding for staff from the Legislature and they had not. Moreover, the pandemic has been over for several years at this point, so we were skeptical of the Department's claims that staffing was the primary problem. When we interviewed Department inspection staff, they told us that the Department's policies on inspections were somewhat confusing and not always consistent. The Department's failure to develop strong inspection policies is likely a large part of the problem. Department management argued that just because it's policies weren't perfect doesn't mean they weren't functional. They pointed out that during fiscal year 25 they were operating with only 25 inspectors, compared to an average of 40 inspectors in previous years. But since they didn't ask for funding, we think this issue is still the Department's fault.

Results

The Department of Food Inspections did not complete a significant portion of required restaurant inspections on time in fiscal year 2025, which may increase risks to public health.

Specifically, the Department completed only about 70% of inspections within the expected timeframe. Department staff attributed these delays primarily to a backlog stemming from the COVID-19 pandemic and noted that the number of inspectors remained below pre-pandemic levels—25 inspectors compared to an average of 40 in previous years. However, the Department did not request additional funding from the Legislature to address staffing shortages. In interviews, inspection staff also pointed to confusion caused by inconsistent or unclear inspection policies, which may have further contributed to the delays. While Department leadership emphasized that its policies were functional and that staff shortages played a significant role, the lack of proactive efforts to secure resources or clarify procedures suggests opportunities for stronger management oversight.