To evaluate the effectiveness of large language models (LLMs) in Adverse Outcome Pathway (AOP) construction, we used GPT-4 for automated annotation to reconstruct five well-documented AOPs described in the AOPwiki. These AOPs are:

1. (AOP38): Protein Alkylation leading to Liver Fibrosis
2. (AOP25): Aromatase inhibition leading to reproductive dysfunction
3. (AOP54): Inhibition of Na+/I- symporter leads to learning and memory impairment
4. (AOP155): Deiodinase 2 inhibition leading to increased mortality
5. (AOP220): Cyp2E1 Activation Leading to Liver Cancer

The general procedure for AOPs reconstruction and verification process is outlined below:

**a. AOP Reconstruction:**

1. Identify MIE and AO: Determine the Molecular Initiating Event (MIE) and the Adverse Outcome (AO) of interest from the AOPwiki or other relevant sources.
2. Input specific questions: Provide GPT-4 with detailed questions about the MIE and AO. Ask GPT-4 to describe the key events (KEs) that connect the MIE to the AO based on available data.
3. Refine model input: Based on the initial output, refine the prompts to ensure that GPT-4 provides comprehensive information about the AOP framework.
4. Integrate annotations: Incorporate the GPT-4-generated annotations into a comprehensive AOP repository. Ensure that each annotated component is correctly linked to its corresponding MIE, KE, and AO.
5. Visualization: Use the annotated data to develop a visual representation of the AOP to illustrate the sequence of events from the MIE through KEs to the AO.

**b. AOP Verification:**

1. Retrieve AOP information: Retrieve original AOP information from the AOPwiki, including clear descriptions of the MIE, KEs, and the AO. Also, retrieve associated detailed annotations and references.
2. AOPs comparison: Compare the GPT-4 reconstructed AOP with the original AOP from AOPwiki. Check if GPT-4 correctly identified all the components and their relationships. Determine if any key components or relationships were omitted or supplemented.
3. Qualitative analysis: Evaluate GPT-4's ability to replicate the original AOP structure. Assess the quality of the annotations in terms of clarity, depth, and relevance. Consider how well GPT-4 captured the biological complexity and the evidence supporting each step in the pathway.

Detailed procedure for AOPwiki: AOP38 (Protein Alkylation leading to Liver Fibrosis) reconstruction were outlined below:

1. Input the specific question:

“Construct the potential Adverse Outcome Pathways (AOPs) using the following keywords: Protein Alkylation, Liver Fibrosis”;

1. GPT-4's annotations:
2. Molecular Initiating Event (MIE): Protein Alkylation

Protein alkylation refers to the covalent attachment of alkyl groups to proteins. This can occur due to exposure to alkylating agents such as certain chemicals or drugs. The alkylation of critical proteins, such as enzymes or structural proteins, can lead to a loss of function or gain of aberrant function.

1. KE1: Altered Protein Function

Alkylated proteins may have altered functions, leading to dysregulation of cellular processes. This can include impaired enzyme activity, disruption of cell signaling pathways, and altered protein degradation processes.

1. KE2: Oxidative Stress and Inflammation

The dysregulation of cellular processes due to protein alkylation can result in increased production of reactive oxygen species (ROS) and the activation of inflammatory pathways. Oxidative stress and inflammation are critical in liver damage and fibrosis.

1. KE3: Hepatocyte Injury and Death

Persistent oxidative stress and inflammation can lead to hepatocyte injury and death. Damaged hepatocytes release pro-inflammatory cytokines and other signals that exacerbate inflammation and tissue damage in the liver.

1. KE4: Stellate Cell Activation

Hepatocyte injury and the subsequent inflammatory response lead to the activation of hepatic stellate cells. These cells play a central role in liver fibrosis by producing excess extracellular matrix (ECM) components, such as collagen.

1. KE5: Extracellular Matrix Deposition

Activated stellate cells secrete ECM proteins, leading to excessive collagen deposition in the liver. This disrupts the normal architecture and function of the liver.

1. Adverse Outcome (AO): Liver Fibrosis

Chronic ECM deposition and the resulting disruption of liver architecture lead to liver fibrosis. Fibrosis impairs liver function and can progress to cirrhosis, liver failure, and increased risk of liver cancer.

Recommended AOP framework:

1. Protein Alkylation (MIE) → 2. Altered Protein Function (KE) → 3. Oxidative Stress and Inflammation (KE) → 4. Hepatocyte Injury and Death (KE) → 5. Stellate Cell Activation (KE) → 6. Extracellular Matrix Deposition (KE) → 7. Liver Fibrosis (AO)

1. Refine the specific question:

“Construct the potential Adverse Outcome Pathways (AOPs) using the following keywords: taking the Protein Alkylation as the Molecular Initiating Event and Liver Fibrosis as the Adverse Outcome”:

1. GPT-4's annotations:
2. Molecular Initiating Event (MIE): Protein Alkylation

Protein alkylation occurs when an alkyl group covalently binds to a protein, typically due to exposure to alkylating agents. This event triggers a series of downstream cellular and molecular events leading to liver fibrosis.

1. KE1: Oxidative Stress

Protein alkylation can lead to an imbalance in the cellular redox state, resulting in the overproduction of reactive oxygen species (ROS). This oxidative stress causes damage to cellular components, such as lipids, DNA, and proteins, which contributes to tissue injury.

1. KE2: Activation of Pro-inflammatory Pathways

Oxidative stress and cellular damage activate pro-inflammatory signaling pathways (e.g., NF-κB, MAPK), leading to the production and release of inflammatory cytokines (e.g., TNF-α, IL-6). These cytokines promote a chronic inflammatory response in the liver.

1. KE3: Hepatocyte Injury and Death

Sustained oxidative stress and inflammation lead to hepatocyte injury and death through necrosis or apoptosis. Dying hepatocytes release damage-associated molecular patterns (DAMPs), which exacerbate the inflammatory response and further tissue damage.

1. KE4: Hepatic Stellate Cell Activation

The inflammatory milieu and signals from damaged hepatocytes, including growth factors like transforming growth factor-beta (TGF-β), activate hepatic stellate cells (HSCs). Activated HSCs differentiate into myofibroblast-like cells that are responsible for producing extracellular matrix (ECM) components.

1. KE5: Excessive Extracellular Matrix (ECM) Deposition

Activated hepatic stellate cells secrete excessive amounts of ECM proteins, primarily collagen, leading to the accumulation of fibrotic tissue. This disrupts the normal architecture and function of the liver, contributing to the development of liver fibrosis.

1. Adverse Outcome (AO): Liver Fibrosis

The chronic and excessive deposition of ECM results in liver fibrosis, characterized by the thickening and scarring of liver tissue. This condition can progress to cirrhosis, liver failure, and an increased risk of liver cancer.

Recommended AOP framework:

Protein Alkylation (MIE) → 2. Oxidative Stress (KE) → 3. Activation of Pro-inflammatory Pathways (KE) → 4. Hepatocyte Injury and Death (KE) → 5. Hepatic Stellate Cell Activation (KE) → 6. Excessive ECM Deposition (KE) → 7. Liver Fibrosis (AO)

1. Refine the specific question:

“Construct the potential Adverse Outcome Pathways (AOPs) using the following keywords: taking the Protein Alkylation as the Molecular Initiating Event and Liver Fibrosis as the Adverse Outcome”:

1. GPT-4's annotations:
2. “Construct the potential Adverse Outcome Pathways (AOPs) using the following keywords: Protein Alkylation as the Molecular Initiating Event (MIE), Liver Fibrosis as the Adverse Outcome (AO)”;
3. “Construct the potential Adverse Outcome Pathways (AOPs) using the following keywords: Protein Alkylation (MIE), Liver Fibrosis (AO)”.

AOP25: Aromatase inhibition leading to reproductive dysfunction

AOP54: Inhibition of Na+/I- symporter leads to learning and memory impairment

AOP155: Deiodinase 2 inhibition leading to increased mortality

AOP220: Cyp2E1 Activation Leading to Liver Cancer