def create\_chatcompletion\_messages(row):

return [

{

"role": "system",

"content": "As a Biological Experiment Standardization Specialist, your task is to standardize the formatting of outcome annotations related to treatment response, survival, and prognosis, integrating them into the single 'Outcome' field."

},

{

"role": "user",

"content": f"Carefully review the info below: {row['Outcome']}\n\n"

"Instructions:\n"

"Before starting the standardization, check the total number of Original Terms provided. Ensure that the order of terms in your Standardized terms matches the exact order in the input data. Standardize each Original Term sequentially without omission, and verify that the total number of Standardized Terms matches the count in the input.\n\n"

"1. Use the definitions and rules for 'Outcome'—which can include up to three components (Response, Survival, Prognosis)—to standardize each term.\n"

"2. Response (for patient or cell line):\n"

" - Patient: Responder, Partial Responder, Stable Disease, Non-Responder, Unknown.\n"

" - Cell line: Sensitive, Partially Sensitive, Resistant, Unknown.\n\n"

"3. Survival (patient only):\n"

" - '[Survival\_Time]: [Status]', e.g. '36 Months: Alive', '12 Months: Deceased'.\n"

" - If no survival data or if it's a cell line: 'NA'.\n\n"

"4. Prognosis:\n"

" - Good Prognosis, Poor Prognosis, Unknown.\n\n"

"5. Combine multiple components with semicolons, e.g. 'Responder; 24 Months: Alive; Good Prognosis'.\n"

"6. If no relevant data is provided, use 'Unknown' or 'NA' as needed.\n\n"

"Format the Output as a Table:\n"

"Create a table with two columns:\n"

"Column 1: List each original term separately.\n"

"Column 2: Provide the standardized term corresponding to each original term.\n\n"

"Ensure accuracy and consistency based on biological standards and domain expertise.\n"

"Do not include any explanations in any format."

}

]