My first task was to test Oracle's capabilities to store and process XML. I used their bulk Data Loader program to load the sample xml NAACCR version 140 file into a CLOB (Character Large Object Blob) field that could then be parsed with Oracles XML functions. The field size for a CLOB can be up to 128 Terabytes. Didn’t take me long to say forget this Oracle XML code given the CDC DLL’s capabilities to parse a NAACCR XML file.

Next task was to create a XMAL application, (Screen pic in attachment), with a back-end of C# for parsing an XML file using the CDC’s XMLPlus.DLL. I got the application to read the XML file and assemble patient/tumor records in a record/field delimited file suitable for bulk loading into our DB. I am adopting what I am calling the KB delimiter in honor of Kathleen B., &#187. (»)

Next Queued tasks

• Integration of the CDC’s EDIT50.dll for scoring and reporting.

• Identify/create better testing data. The Current XML sample file is a NAACCR version 140 ‘A’ record type. I need to test 160 and 180 versions for record type I.

• Continue refining/testing load to add error handling and reporting.

XAML App:

