Figure 3.5: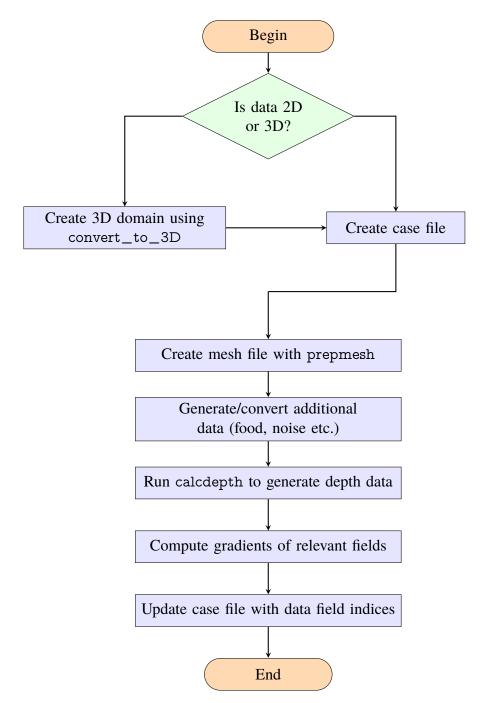 Flowchart showing data preparation steps