Model performance on in-domain RT test dataset (any SDoH)