Key scores per overlap threshold on validation set