Supplementary Materials

A. Experimental Details

Table 1. Dataset details

	CT-MRI	FLAIR-T1
#total	219	1470
#train	197	1251
#test	22	219
shape	160×160×160	176×176×160
voxel size	$1\times1\times1mm^3$	

Table 2. Implementation details

	CT→MRI	FLAIR→T1
batch size	16	8
#iterations	100,000	150,000
λ	0.5	0.05
time step	1000	
sampling step	100	
GPU	NVIDIA RTX A6000	

B. Additional Qualitative Comparison

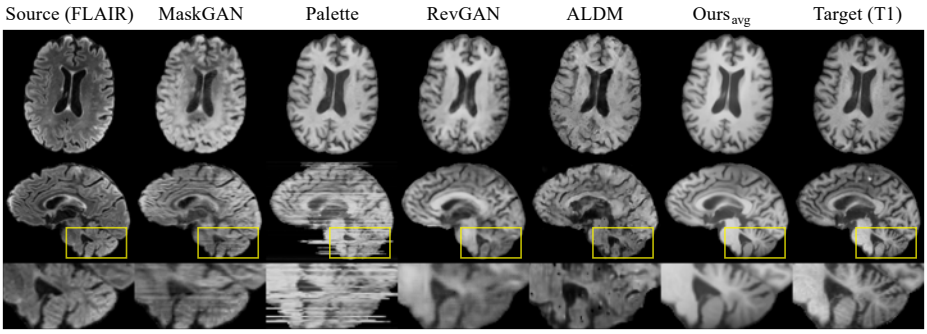

Fig. 1. Qualitative comparison with baselines (FLAIR→T1)

C. Derivation of Score Function $\boldsymbol{S}_\theta(\boldsymbol{X}_t, \boldsymbol{c}_{SKC}, t)$ for θ Trained as $\theta = \theta^*$

Since $\boldsymbol{E}_{\theta^*,t} = CP_{\theta^*}(\boldsymbol{X}_t, \boldsymbol{c}_{SKC}, t) = m_t(\boldsymbol{Y} - \boldsymbol{X}_0) + \sqrt{\delta_t}\boldsymbol{\epsilon}_t$, where $\boldsymbol{\epsilon}_t \in \mathbb{R}^{Z\times H\times W}$,

$CP_{\theta^*}(\boldsymbol{X}_t, \boldsymbol{c}_{SKC}, t) = \boldsymbol{X}_t - \boldsymbol{X}_0$ ($\because \boldsymbol{X}_t = (1-m_t)\boldsymbol{X}_0 + m_t\boldsymbol{Y} + \sqrt{\delta_t}\boldsymbol{\epsilon}_t$).

Then, $CP_{\theta^*}(\boldsymbol{X}_t, \boldsymbol{c}_{SKC}, t) = m_t[\boldsymbol{Y} - \{\boldsymbol{X}_t - CP_{\theta^*}(\boldsymbol{X}_t, \boldsymbol{c}_{SKC}, t)\}] + \sqrt{\delta_t}\boldsymbol{\epsilon}_t$

$\iff \boldsymbol{\epsilon}_t = \frac{1}{\sqrt{\delta_t}}\{m_t(\boldsymbol{X}_t - \boldsymbol{Y}) + (1-m_t)CP_{\theta^*}(\boldsymbol{X}_t, \boldsymbol{c}_{SKC}, t)\}$.

Thus, $\boldsymbol{S}_{\theta^*}(\boldsymbol{X}_t, \boldsymbol{c}_{SKC}, t) = \nabla_{\boldsymbol{X}_t}\log q_{BB}(\boldsymbol{X}_t \mid (1-m_t)\boldsymbol{X}_0 + m_t\boldsymbol{Y}, \boldsymbol{c}_{SKC})$

$= -\frac{[\boldsymbol{X}_t - \{(1-m_t)\boldsymbol{X}_0 + m_t\boldsymbol{Y}\}]}{\delta_t} = -\frac{\boldsymbol{\epsilon}_t}{\sqrt{\delta_t}}$ (cf. $\nabla_{\tilde{\boldsymbol{x}}}\log q_\sigma(\tilde{\boldsymbol{x}} \mid \boldsymbol{x}) = -\frac{\tilde{\boldsymbol{x}} - \boldsymbol{x}}{\sigma^2}$ in [1]),

since $\boldsymbol{X}_t$ is the Gaussian perturbation of $(1-m_t)\boldsymbol{X}_0 + m_t\boldsymbol{Y}$.

$\therefore \boldsymbol{S}_{\theta^*}(\boldsymbol{X}_t, \boldsymbol{c}_{SKC}, t) = -\frac{1}{\delta_t}\{m_t(\boldsymbol{X}_t - \boldsymbol{Y}) + (1-m_t)CP_{\theta^*}(\boldsymbol{X}_t, \boldsymbol{c}_{SKC}, t)\}$.

D. Qualitative Results of Ablation Study

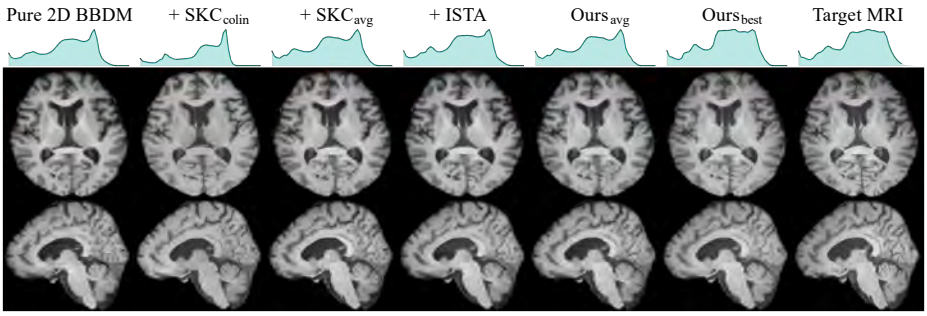

Fig. 2. Qualitative comparison with baselines (CT→MRI)

E. Additional Study for Style Key Conditioning

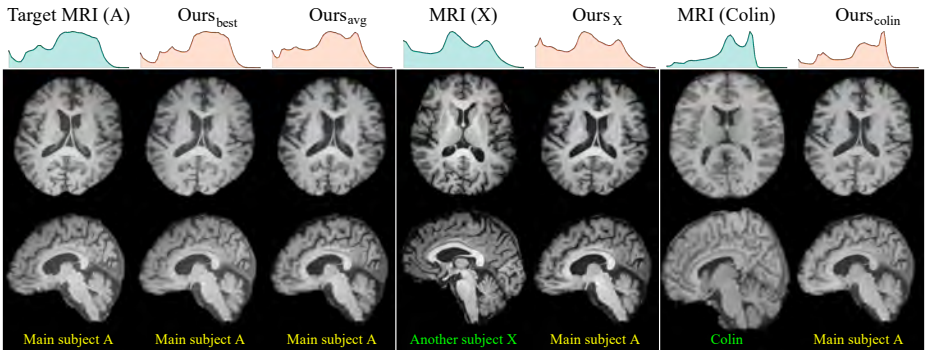

Fig. 3. Qualitative results for various style keys (CT→MRI). Figure shows real MRIs of main subject A, another subject X, and Colin 27 Average Brain (public access), along with synthetic MRIs of subject A. Both Ours$_{X}$ and Ours$_{colin}$ present synthetic MRIs of subject A, each transformed into the respective styles of X and Colin. From the analysis of the intensity histogram presented above, we observe that Ours$_{X}$ closely matches the MRI histogram of subject X, while Ours$_{colin}$ aligns with that of Colin. This demonstrates that the style of MRIs can be effectively controlled using SKC.

References

1. Song, Y., Ermon, S.: Generative modeling by estimating gradients of the data distribution. Advances in neural information processing systems **32** (2019)