Structural Connectome Analysis using a Graph-based Deep Model for Prediction of Non-Imaging Variables

Anees Kazi,^{1,2*} Jocelyn Mora,^{1} Bruce Fischl,^{1,2} Adrian V. Dalca,^{1,2,3} Iman Aganj^{1,2}

^{1}Athinoula A. Martinos Center for Biomedical Imaging, Radiology Department,

Massachusetts General Hospital, Boston, USA

^{2}Radiology Department, Harvard Medical School, Boston, USA

^{3}CSAIL, Massachusetts Institute of Technology, Cambridge, USA

*Correspondence: AKAZI1@mgh.harvard.edu

April 9, 2026

Keywords: Graph neural network (GNN), Graph convolutional network (GCN), Brain connectivity, Diffusion MRI, Prediction, Deep learning, Age, Dementia

Abstract: We address the prediction of non-imaging variables based on structural brain connectivity derived from diffusion magnetic resonance images, using graph-based machine learning. We predict age and the mini-mental state examination (MMSE) score as examples of a demographic and a clinical variable. We propose a machine-learning model inspired by graph convolutional networks (GCNs), which takes a brain connectivity graph as input and processes the data separately through a parallel GCN mechanism with multiple branches. The novelty of our work lies in the model architecture, especially the Connectivity Attention Block, which learns an embedding representation of brain graphs while providing graph-level attention. We show experiments on publicly available datasets of PREVENT-AD and OASIS3. We validate our model by comparing it to existing methods and via ablations. This quantifies the degree to which the connectome varies depending on the task, which is important for improving our understanding of health and disease across the population. The proposed model generally demonstrates higher performance especially for age prediction compared to the existing machine-learning algorithms we tested, including classical methods and (graph and non-graph) deep learning.

1 Keywords:

Graph neural network (GNN), Graph convolutional network (GCN), Brain connectivity, Diffusion MRI, Prediction, Deep learning, Age, Dementia

2 Introduction

A structural connectome is a comprehensive map of the physical (anatomical) connections between different regions of the brain. It represents the white matter pathways that link various brain regions, forming a network that supports communication and information flow across the brain (Hofman, 2015; Sporns et al., 2005). Analyzing the structural connectome provides insight into neural circuitry, connectivity patterns, and their implications for cognition and behavior, contributing to our understanding of brain development, aging, and the impact of neurological conditions on brain wiring (Hagmann et al., 2008; F. Zhang et al., 2022). Connectomic analysis is a relatively new avenue for studying the brain, which, while not yet included in clinical practice, holds promise for neuroscientific discoveries and biomarker development. Given the complexity and the vast amount of data involved, deep learning (DL) techniques are likely to be effective in analyzing the connectome. Structural brain connectivity analysis is a data-driven approach that explores the overall structural connections in the brain. It involves studying patterns of anatomical connectivity via fiber tracking (tractography) throughout the brain (Zalesky et al., 2012).

Structural connectome graphs have been used to study a wide range of neurological and psychiatric disorders, including Alzheimer's disease (AD) (Aganj et al., 2023; Amoroso et al., 2017; Frau-Pascual et al., 2021; He et al., 2008; J. Wang et al., 2018), schizophrenia (Karlsgodt et al., 2010; Shi et al., 2012; Y.-m. Wang et al., 2020), autism spectrum disorders (ASD) (Tolan & Isik, 2018), and Parkinson's disease (Arrigo et al., 2019; Jellinger, 2022; X. Zhang et al., 2018; Y. Zhang et al., 2019), as well as to understand cognitive decline (X. Xu et al., 2021) and normal brain development and aging (Coelho et al., 2021; Damoiseaux, 2017; Dennis et al., 2013; Lewis et al., 2022; Neudorf et al., 2024). Numerous studies have identified specific structural and functional changes in the brains of patients with AD, such as altered functional activation, decreased volume or thickness of gray matter in certain brain areas, and abnormalities in the connections between different regions (Xie & He, 2012).

Structural brain connectivity analysis using artificial intelligence (AI) is an emerging field (Dubost, 2020; Sjöblom et al., 2020). The exploration of AI applications in deciphering dMRI data to study brain connectivity is an evolving domain that needs further investigation (Faiyaz et al., 2023). In this paper, we develop a technique to analyze dMRI-derived connectivity using DL for age and dementia prediction. Recently, Prescott et al. (2022) have shown that structural connectivity is a crucial factor in identifying early-onset AD risk, as individuals with a genetic predisposition show lower connectivity, especially in the frontoparietal control network, and this reduced connectivity is linked to the estimated time until dementia symptoms emerge. Toward this goal, one of the key tasks we address is predicting the Mini-Mental State Examination (MMSE) score of subjects directly from their structural connectomes using our proposed AI-based framework, thereby enabling quantitative estimation of cognitive decline associated with dementia.

Further, studying brain aging through the structural connectome is crucial as it provides insight into how brain connectivity and network organization change over time, which can be related to cognitive decline, neurodegenerative diseases, and overall brain health (Meunier et al., 2009). Accurate brain age prediction also serves as a potential biomarker for early detection of atypical aging trajectories, helping to identify individuals at risk of conditions such as AD and other forms of dementia. The task of predicting clinical and demographic variables from structural connectomic data, however, poses significant challenges due to the high dimensionality and non-Euclidean nature of brain connectivity. A recent work (Joo et al., 2023) shows that convolutional neural networks show suboptimal performance when used standalone, compared to using them in combination with Multi Layered Perceptrons (MLPs). Traditional DL approaches for age prediction have predominantly utilized convolutional neural networks (CNNs) applied to structural MRI data. A recent work (Peng et al., 2023) proposed a DL algorithm that leverages brain structural imaging data and enhances prediction accuracy by integrating biological sex information, employing a dataset of 3004 healthy subjects aged 18 and above using CNN algorithms on T1-weighted images. Similarly, lightweight CNN architectures such as Simple Fully Convolutional Networks (SFCN) have demonstrated accurate disease prediction with neuroimaging data, with CNNs showing the ability to accurately predict chronological age in healthy individuals from structural MRI brain scans (Cole et al., 2022; Peng et al., 2021). Machine learning (ML) algorithms trained to estimate age from brain structural MRI using DL-based approaches have revealed genetic associations, with new DL methods uncovering associated sequence variants (Cole et al., 2019). Recent advances include anatomical feature attention-enhanced 3D-CNNs for age prediction from structural MRI data, with CNN architectures capturing structural feature changes of brain aging based on MRI to predict age in healthy individuals (Feng et al., 2024; Zhao et al., 2024).

DL approaches for MMSE prediction have been less extensively explored compared to age prediction, with most studies focusing on joint prediction tasks combining disease diagnosis and cognitive assessment. Ryu et al. (2019) developed a 3D CNN architecture for both AD classification and joint MMSE score prediction using resting-state functional MRI scans of 331 participants, obtaining functional 3-dimensional independent component spatial maps as features for both classification and regression tasks (Ryu et al., 2019). The limited number of studies specifically targeting MMSE prediction using DL suggests this remains an underexplored area, with most cognitive assessment predictions relying on multimodal approaches or functional connectivity data rather than structural MRI alone.

Graph neural networks (GNNs), particularly the graph convolutional network (GCN) architecture (S. Zhang et al., 2019), provide a powerful and flexible framework for analyzing brain connectivity data.

GCNs operate on graph representations of the brain, where nodes correspond to anatomical regions and edges encode structural connectivity (e.g., fiber density or tract strength), enabling the model to integrate information from each region along with the context provided by its connected neighbors. Thus, GCNs leverage the rich structural information in the connectome to make accurate predictions. They do so by performing iterative message-passing between neighboring nodes in the graph, using learnable functions to aggregate and transform information from neighboring nodes, and updating the features of each node based on the aggregated information. This allows GCNs to capture the complex relationships between brain regions and their connections, and make predictions based on this information. A GCN-based model is promising for analyzing the structural connectome and has shown great potential for improving our understanding of neurological and psychiatric disorders (Kazi, Shekarforoush, Arvind Krishna, et al., 2019; Kazi, Shekarforoush, Kortuem, et al., 2019), as well as normal brain development and aging (X. Li et al., 2020).

Recent GCNs have been employed for functional brain network analysis (X. Li et al., 2021), psychiatric disorder diagnosis (Zheng et al., 2024), optimization of GNN architectures for schizophrenia spectrum disorder prediction (S. Wang et al., 2024), classification of ASD versus hyper complex brain networks (Hu et al., 2021), and sex classification (Ktena et al., 2018). Many papers have shown work using structural brain networks for causal inference (Wein et al., 2021), brain age estimation (Moon et al., 2024), early diagnosis of AD (Y. Zhang, He, et al., 2023), and schizophrenia diagnosis (Sebenius et al., 2021), achieving high performance compared to existing ML, including DL, methods. GCNs have also been combined with recurrent neural networks to predict sex on temporal functional brain connectivity graphs (Kazi et al., 2022). A spectral GCN has been employed for region-of-interest identification in functional connectivity graphs and for sex classification (Arslan et al., 2018). In a multi-modal integration technique, Qu et al., 2025, the authors proposed a MaskGNN framework that integrates multimodal brain connectivity data for predictive analysis. Specifically, functional connectivity (FC) from fMRI and structural connectivity (SC) from DTI were combined at the node level and used as inputs to the MaskGNN model. In the latent space, the model further incorporated anatomical statistics (AS) from sMRI and computed structure–function coupling to capture complementary relationships between modalities. The aggregated features were processed through MaskGNN embedding, pooling, and readout layers, enabling effective multimodal feature learning.

Methodologically, most GCN methods use the Laplacian-filter based spectral implementation (Kipf & Welling, 2016). GCNs suffer from limited generalization to new graphs due to their dependency on the fixed graph Laplacian (Q. Li et al., 2018; Oono & Suzuki, 2019), and they struggle with over-smoothing in deep architectures where node representations become indistinguishable. Graph attention networks (GATs) (Veličković et al., 2017) incorporate attention mechanisms for adaptive edge weighting, enabling nodes to selectively attend to their neighborhood features without requiring costly matrix operations or prior knowledge of graph structure. GATs leverage masked self-attentional layers to address shortcomings of prior graph convolution methods by enabling nodes to attend over their neighborhoods' features and implicitly specify different weights to different nodes in a neighborhood without requiring costly matrix operations or depending on knowing the graph structure upfront. Despite these advantages, GATs are computationally inefficient for capturing long-range dependencies due to over-smoothing issues (Dasoulas et al., 2021) and exhibit high computational costs when applied to dense graphs (Wu et al., 2020).

Dynamic Graph CNNs (DGCNNs) (Phan et al., 2018) are employed for temporal graph analysis of multi-level graph structures and enable hierarchical learning through spatiotemporal integration of spatial and temporal data. While DGCNNs can adapt to changing graph topologies over time, they suffer from computational inefficiency and instability due to repeated graph construction at each layer, leading to high memory usage and potential noise amplification. Both GCN and GAT methods face scalability challenges with large graphs (S. Liu et al., 2020), requiring specialized sampling or approximation techniques to maintain computational tractability in real-world applications.

In this paper, we propose a novel graph-based DL model to improve the prediction of demographic and clinical variables from structural brain connectivity. We introduce a computationally efficient and robust architecture that combines three key components: 1) GCNs with residual connections (ResGCN (Pei et al., 2022)) to support brain graph data processing and analysis, 2) fully connected (FC) layers followed by non-linearity for efficient structure-independent feature learning, and 3) a novel Connectivity Attention Block (CAB) that dynamically prioritizes informative inter-regional brain connections. Attention is widely used in recent methods, is competitive with convolutions (Bahdanau, 2014; Vaswani, 2017), and is used in brain image analysis (Jiao et al., 2023; Ranjbarzadeh et al., 2021; Y. Zhang, Teng, et al., 2023). Therefore, the design involving attention allows the model to capture rich representations at both the node (brain region) and subject levels, bridging fine-grained neuroanatomical patterns and global connectivity profiles. Following our previous work demonstrating the improvement that structural connectivity data brings to cortical registration (Zhou et al., 2026) and segmentation (Łabiak et al., 2026), here we show how it can help in prediction. Experimental results on two publicly available neuroimaging datasets show that our method outperforms both traditional ML approaches and state-of-the-art graph-based DL models. A preliminary conference version of the proposed network architecture (excluding the new CAB) for the task of sex classification has been previously presented (Kazi, Mora, et al., 2023).

The architecture introduces a combination of complementary components that collectively overcome these individual weaknesses of traditional methods. The integration of ResGCNs with residual connections (Pei et al., 2022) is expected to tackle the over-smoothing problem that plagues standard GCNs in deep architectures, enabling the model to retain discriminative power across multiple layers. The inclusion of FC layers provides structure-independent feature learning capabilities, compensating for the limited generalization issues inherent in graph Laplacian-dependent methods by allowing the model to capture global patterns beyond local graph topology. Most importantly, the novel CAB represents a departure from traditional uniform message aggregation to adaptive, data-driven connection weighting, directly addressing the challenge of dense brain connectivity networks where standard GNN methods exhibit performance degradation. This attention-based approach aligns with recent advances in topological cycle graph attention networks that delineate functional backbones within brain functional graphs by distinguishing key pathways essential for signal transmission from non-essential, redundant connections. By combining these three components, the architecture creates a robust framework that maintains computational efficiency while capturing both fine-grained neuroanatomical patterns and global connectivity profiles, effectively bridging the gap between the theoretical capabilities of GNNs and the practical demands of connectomic analysis for age and cognitive-assessment prediction.

The remainder of this paper, which provides a comprehensive evaluation of the proposed architecture and its implications for neuroimaging research, is organized as follows. We begin with a detailed presentation of the model architecture in the Method section, including the mathematical formulations of all the model components, the design specifications of the FC layers, and the new CAB mechanism. Subsequently, we describe the experimental design in the Experiments section, encompassing dataset descriptions, preprocessing protocols, evaluation metrics, implementation details, and baseline comparisons with both traditional ML approaches and state-of-the-art graph-based DL models for cross-validation and test data. The Results and Discussion section presents quantitative performance analyses across multiple neuroimaging tasks, including age and MMSE score prediction, along with ablation studies demonstrating the contribution of each architectural component. Following the empirical evaluation, we provide a thorough discussion of the findings, addressing the model's strengths, limitations, and computational efficiency compared to existing methods. Finally, we explore the clinical implications of our approach, discussing how the improved prediction accuracy and interpretability features of the attention-based connectivity analysis can inform clinical decision-making processes and contribute to a better understanding of brain aging and patterns of cognitive decline in neurological and psychiatric disorders.

3 Method

Given a dataset of S subjects, each subject has an associated brain graph $\mathbf{G} \in \mathbb{R}^{N\times N}$ with N denoting the number of brain regions (graph nodes), a feature matrix $\mathbf{X} \in \mathbb{R}^{N\times M}$ with M the number of features

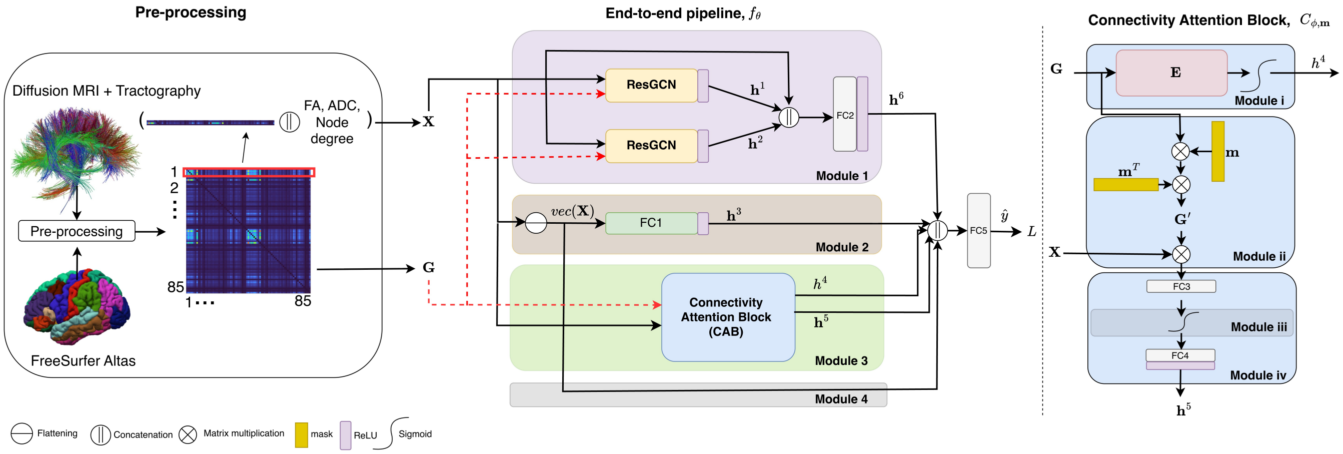

Pre-processing
Diffusion MRI + Tractography
FA, ADC, Node degree
Pre-processing
1
2
85
1
85
FreeSurfer Altas
X
G
End-to-end pipeline, f_θ
ResGCN
ResGCN
$\mathbf{h}^1$
$\mathbf{h}^2$
FC2
$\mathbf{h}^6$
Module 1
$vec(\mathbf{X})$
FC1
$\mathbf{h}^3$
Module 2
Connectivity Attention Block (CAB)
h^4
$\mathbf{h}^5$
Module 3
Module 4
FC5
$\hat{y}$
L
Connectivity Attention Block, $C_{\phi,\mathbf{m}}$
G
E
h^4
Module i
$\mathbf{m}$
$\mathbf{m}^T$
$\mathbf{G}'$
X
Module ii
FC3
Module iii
FC4
Module iv
$\mathbf{h}^5$
Flattening
Concatenation
Matrix multiplication
mask
ReLU
Sigmoid

Figure 1: Overview of an end-to-end DL architecture for brain connectivity analysis using dMRI and tractography. The pre-processing module (left) includes the dMRI-based tracts forming the structural connectome matrices. dMRI-derived features (FA, ADC, degree) are concatenated with each row of the connectivity matrix and then used as node features. The model (middle) processes input features and graphs through residual GCNs, fully connected (FC) layers, and a novel Connectivity Attention Block (CAB, right), which refines the graph structure and highlights key connections. Outputs from all modules are combined to predict phenotypes. See the Method section for more details.

per node, and a label $y \in \mathbb{R}$ (i.e., value of the clinical or demographic variable to be predicted). The task is to predict y, for which we define a model f_θ as:

$$\hat{y} = f_\theta(\mathbf{G}, \mathbf{X}), \tag{1}$$

where θ is the set of learnable parameters. The model we propose comprises four sub-parts, modules 1 to 4, as shown in Figure 1. Each module is designed to process the combination of $\mathbf{X}$ and $\mathbf{G}$ to produce a latent embedding. The last part combines all the previous outputs to produce the predicted label $\hat{y}$. We describe all four modules separately and then explain the whole end-to-end model.

3.1 Module 1: Graph Convolution

This module comprises two branches, each with a graph convolution with a different embedding size, and a skip connection as shown in Figure 1. We employ a combination of ResGCN layers (Bresson & Laurent, 2017), which captures low-level features of the graph. Each ResGCN (with a different embedding size) transforms the same input data into a different space, enabling the model to learn diverse representations of the graph data. One layer focuses on capturing low-dimensional, essential features, while the other learns higher-dimensional, more nuanced relationships. The skip connection is then combined with these learned representations, allowing the model to leverage both the basic, essential features and the more intricate relationships within the graph. The outputs of these layers are fed to an FC layer (shown as FC2 in Figure 1) to produce the output of the module, $\mathbf{h^6}$.

3.2 Module 2: Fully Connected (FC1) Layer

In this layer, we remove any structure that is present in the feature vector sequence **X** and simply extract information from the raw data. This layer can be mathematically defined as: $\mathbf{h^3} = \rho\left(FC1(vec(\mathbf{X}))\right)$, where $vec(\mathbf{X}) \in \mathbb{R}^{NM \times 1}$ is the flattened matrix **X**, and ρ is a non-linearity (ReLU). An FC layer processing node in parallel with GCN layers offers several potential benefits. It allows the model to learn an independent representation of the node features, capturing information that might not be directly reflected in the graph structure. This is particularly advantageous when the node features themselves hold significant information for the task at hand.

3.3 Module 3: Connectivity Attention Block (CAB)

We propose a customized CAB, $C_{\phi,\mathbf{m}}$, so that the model learns an embedding representation of a brain graph, as well as to provide a graph-level attention mechanism. The CAB, depicted in Figure 1 (right), is represented as

$$h^4, \mathbf{h^5} = C_{\phi,\mathbf{m}}(\mathbf{G}, \mathbf{X}), \tag{2}$$

where h^4 is a scalar embedding of each subject's brain connectivity, which is also treated as a factor of importance of each subject with respect to the population. $\mathbf{h^5}$ is a lower-dimensional representation that is sculpted out of $\mathbf{G}$ and $\mathbf{X}$. The external trainable parameter $\mathbf{m}$, initialized uniformly, is the attention mask assigned to the nodes (i.e. regions). All are defined as:

$$C_{\phi,\mathbf{m}} = \begin{cases} h^4 = \sigma\left[\langle \mathbf{G}, \mathbf{E} \rangle_F\right] \\ \mathbf{h^5} = \rho[f_4(\sigma[f_3(\mathbf{G}\mathbf{m}\mathbf{m}^T\mathbf{X})])] \end{cases} \tag{3}$$

h^4 is the Frobenius inner product of the **G** matrix by a weight matrix $\mathbf{E} \in \phi$ learned by the model, followed by sigmoid non-linearity, σ. For $\mathbf{h^5}$, we effectively project **G** onto a single one-dimensional orientation within the N-dimensional space, from which we create a rank-one matrix $\mathbf{G}' := \mathbf{G}\mathbf{m}\mathbf{m}^T$, with the projection weights $\mathbf{m}$ learned by the model. f_3 and f_4 are the FC layers $FC3$ and $FC4$ as shown in Figure 1, both followed by non-linearities. Here, instead of applying attention to each element of $\mathbf{G}$, we leverage the matrix properties of $\mathbf{G}$. Our goal is to retain a single representative rank-one matrix $\mathbf{G}'$ that best helps the prediction task at hand. This operation simplifies the representation of the matrix while preserving its most significant features, along the same lines as singular-value decomposition. The vector $\mathbf{m}$ is optimized jointly with the rest of the network to maximize prediction performance, by weighting the region-wise features by $\mathbf{m}$ and projecting the results onto the direction $\mathbf{Gm}$ (i.e., eigenvector of $\mathbf{G}'$), effectively narrowing the attention to a single orientation in the N-dimensional space.

Importantly, CAB is not intended to serve as the sole source of model expressivity. Higher-order graph structure and complex interaction patterns are captured by the residual GCN blocks and the fully connected branches (Modules 1 and 2), while CAB provides a complementary, global, and interpretable reweighting of brain regions and their associated connections. This design allows CAB to enhance the model's representational capacity without replacing the richer structural modeling performed by the other components.

CAB works directly on the full adjacency matrix G and learns a single dominant eigen-direction-like m (a region mask) to produce a task-adaptive rank-1 approximation $\tilde{G} = Gmm^T$, distilling global structural connectivity into a low-dimensional, interpretable embedding. This allows CAB to capture global patterns that local/node-level attentions cannot. To compare with existing methods: GAT (Veličković et al., 2017) computes attention only between neighboring nodes (local 1-hop edges), whereas CAB derives the attention weights from the entire connectivity matrix. Low-rank attention (S. Wang et al., 2020) methods approximate sequence attention via query-key projections. CAB instead leverages the symmetry of brain connectivity matrices via a rank-1 projection, avoiding unnecessary decompositions. Spectral attention (N. Liu et al., 2024) methods rely on fixed Laplacian filters. CAB learns a data-driven eigen-like vector end-to-end, optimized directly for the prediction task.

3.4 Module 4: Skip Connection

We add an overall skip connection (Module 4 in Figure 1) to help mitigate the vanishing gradient problem and allow effective information propagation across layers by directly connecting $vec(\mathbf{X})$ to the final layer.

3.5 Data Fusion and Loss Function

The outputs of all four modules are concatenated as $\mathbf{X_{all}} = [\mathbf{h^3}, h^4, \mathbf{h^5}, \mathbf{h^6}, vec(\mathbf{X})]$ and fed to an FC layer to produce the final prediction $\hat{y} = FC5(\mathbf{X_{all}})$

For the prediction task, we use a customized loss function made of two parts. The first part is the Huber

Table 1: Description of dataset sizes (number of available scans) and partitioning for PREVENT-AD and OASIS3 dataset. Due to missing demographic data, nine subjects were removed from the OASIS3 dataset. Each subject may have multiple samples collected at different time points, hence there are overall more samples than subjects.

Name	Subjects	Total samples	90% Cross-validation set	Unseen 10% test set
PREVENT-AD	347	789	714	75
OASIS3	771	1294	1164	121

loss, $L_\delta(y - \hat{y})$, which is a piecewise continuous function, defined as:

$$L_\delta(e) = \begin{cases} \frac{1}{2}e^2 & \text{if } |e| \leq \delta, \\ s\delta \cdot (|e| - \frac{1}{2}\delta) & \text{otherwise,} \end{cases} \tag{4}$$

where $\delta > 0$ is a threshold that controls the location of the transition between quadratic and linear loss behaviors. This function applies a quadratic loss when the absolute error is less than or equal to δ and a linear loss when the error exceeds this threshold, effectively combining the differentiability of the quadratic loss at small errors with the robustness of the linear loss to outliers. The second term of the loss, L^R, is the regularizer for $\mathbf{m}$ defined as $L^R = \sum_i^N |m_i| - \sum_i^N |m_i| \log |m_i|$ (Kazi, Farghadani, et al., 2023). The overall loss for the model is therefore $L = L_\delta + \beta L^R$, where $\beta > 0$ is a weighting factor.

4 Experiments

4.1 Datasets

We use two publicly available datasets. The details are given below.

Pre-symptomatic Evaluation of Experimental or Novel Treatments for Alzheimer's Disease (PREVENT-AD) (Leoutsakos et al., 2016) is a publicly available dataset that provides a comprehensive set of data on individuals who are at risk for developing AD (https://prevent-alzheimer.net). The database contains neuroimaging studies such as MRI (including dMRI) and PET scans, a range of demographic, clinical, cognitive, and genetic data, as well as data on lifestyle factors such as diet and exercise. The dataset comprises 347 subjects, some with multiple (longitudinal) dMRI scans, totaling 789 dMRI scans.

The third release of **Open Access Series of Imaging Studies (OASIS3)** (LaMontagne et al., 2019) is a longitudinal neuroimaging dataset with clinical and cognitive scores for normal aging and AD, and is provided freely to researchers (http://www.oasis-brains.org). The OASIS3 dataset contains MRI scans (including dMRI), cognitive assessments, demographic information, and clinical diagnoses for subjects, including healthy controls, individuals with mild cognitive impairment, and AD patients. From this dataset we used 1294 brain scans from 771 subjects. Please refer to Table 1 for the overview of the datasets.

4.1.1 Pre-processing

We used FreeSurfer (Fischl, 2012) to process the databases (additionally applying the longitudinal processing pipeline (Reuter et al., 2012) for PREVENT-AD). We then ran the FreeSurfer diffusion processing pipeline and propagated the 85 automatically segmented cortical and subcortical regions from the structural to the diffusion space. These 85 regions act as the nodes in our graph setup. Next, we used the public CSA-ODF and Hough Tractography toolbox (http://www.nitrc.org/projects/csaodf-hough) to reconstruct the diffusion orientation distribution function in constant solid angle (Aganj et al., 2010), run Hough-transform global probabilistic tractography (Aganj et al., 2011) to generate 10,000 optimal streamlines per subject, compute symmetric structural connectivity matrices, and augment the matrices with indirect connections (Aganj et al., 2014). More details on the pipeline can be found in our previous publication (Aganj et al., 2023). Once we had all the graphs G_i, we performed a population-level normalization on edge weights.

For node features, we used the structural connectivity map of the node region to the rest of the regions as well as mean fractional anisotropy (FA), mean apparent diffusion coefficient (ADC), volume, and the degree of the node. In predicting age and dementia, dMRI analysis can benefit from a combination of node features. FA and ADC, reflecting white matter organization and water diffusion, respectively, capture microstructural changes due to aging and AD, whereas mean segmentation volume captures brain region size that is potentially reduced in age-related atrophy or AD. Finally, the node degree (total connectivity to a brain region) reveals alterations in brain network connectivity patterns. Therefore, for each subject, we obtained $\mathbf{G}_i \in \mathbb{R}^{85\times85}$ and corresponding $\mathbf{X}_i \in \mathbb{R}^{85\times89}$ (i.e., $N = 85$ and $M = 89$). The four additional features capture complementary information beyond connectivity and standard diffusion measures, improving robustness and performance. FA captures the degree of directional water diffusion and serves as a microstructural marker of white-matter organization, whereas ADC reflects the overall magnitude of diffusion within tissue. The regional volume feature quantifies morphometric variation across the segmented brain regions, and node degree represents the number of structural connections incident to each node, thereby summarizing its topological centrality within the connectome.

4.1.2 Implementation Details

All the experiments were run via 10-fold cross-validation with the same folds across methods and experiments. For model robustness, we added zero-mean Gaussian noise with a standard deviation of 0.1 to the training samples. All the experiments were run on a Linux machine with 512 GB of RAM, an Intel (R) Xeon (R) Gold 6256 CPU @ 3.60 GHz, and an NVIDIA RTX A6000 (48 GB) graphics processing unit. The total number of parameters used in the proposed model was 5073, which was comparable to GCNConv (2453), DGCNN (4653), GraphConv (4653), ResGatedGraphConv (RGGC) (9128), and GINConv (2683). In our experiments, the output embedding sizes for $\mathbf{h^1}$, $\mathbf{h^2}$, $\mathbf{h^3}$, and $\mathbf{h^5}$ were 25, 20, 5, and 2, respectively. The values of $\beta = 0.01$ chosen empirically in the loss function. Each input feature and edge weight was normalized using a fixed population-level min–max scaling, where the range was computed from the combined training and validation (but not the test) cohorts. This ensures a consistent scale across features and prevents high-variance attributes (e.g., volumetric measures) from disproportionately influencing the model gradients.

4.2 Experimental Design

4.2.1 Evaluation Metrics

Performance is evaluated using four complementary metrics: Root Mean Square Error (RMSE) and Mean Absolute Error (MAE), which measure the magnitude of prediction errors, as well as Pearson's Correlation Coefficient (PC) and Spearman's Correlation Coefficient (SC), which assess the strength and direction of the linear and rank-order relationships, respectively, between predicted and ground-truth values. RMSE is particularly sensitive to large errors, making it useful for penalizing significant deviations, while MAE provides a more interpretable average error magnitude. Pearson's correlation captures linear associations, whereas Spearman's accounts for monotonic but potentially non-linear trends. Together, these metrics offer a comprehensive assessment of both the accuracy and reliability of the model predictions across all three tasks, enabling a nuanced understanding of performance from both error-based and correlation-based perspectives.

4.2.2 Cross-validation and Unseen Test Set Experiments

We divide the entire dataset into a 90:10 split. We set aside the 10 % data as an unknown test set and only use it once the training of the model is done. We use the remaining 90% data for cross-validation by splitting the data again into 90:10 ratio. The dataset split is shown in Table 1. We use `sklearn` (Feurer et al., 2020) to perform 10-fold cross-validation using experiments for six classical ML and 5 DL based methods. All the sets are created based on subject IDs instead of sample IDs to avoid data leakage and ensure that no information from the same subject appears in both the training and validation/test sets. We keep 10% of the data aside from each dataset so as not to heuristically fit the model to the entire data, and then test the model at the end on the unseen data, the results of which are reported in the next section.

4.2.3 Baselines and Comparative Methods

We performed a comparative evaluation of various conventional ML and advanced DL models for the age prediction task across the two datasets of PREVENT-AD (789 scans) and OASIS3 (1294 scans), as well as for MMSE prediction in the OASIS3 dataset, while ensuring that all scans of the same subject appeared only in a single (train, validation, or test) set. As a baseline, we employ a naive model that consistently predicts the mean (resp. median) of the training set. We evaluate its performance using RMSE (resp. MAE) on the validation set. These baseline metrics provide a sanity check, establishing the minimum performance threshold that all other models are expected to surpass. Since this model outputs a constant value, its correlation coefficients (PC and SC) with the ground truth are zero. The conventional ML models that we tested include linear regression (LR), support vector regression (SVR), decision trees, regression trees, ensemble trees, and neural networks (all implemented in the `sklearn` package). We further compared our method with DL approaches such as the multilayer perceptron (MLP), and more advanced GNN models such as GCN, GIN, GraphConv, and ResGCN.

4.2.4 Ablation Studies

We performed ablation experiments to systematically evaluate the individual contributions of each architectural component in the proposed model and to understand their impact on overall predictive performance. By selectively removing key modules—such as the FC1 block, CAB block, skip connection, and GCN block—and observing changes in prediction accuracy, we aimed to identify which components are essential for capturing complex patterns in the data. This approach helps to validate the necessity of each element within the model design, ensuring that the architecture is both effective and interpretable. Moreover, it highlights the role of graph-based learning in enhancing the model's capacity to model relationships within neuroimaging data, thereby justifying the inclusion of the GCN block in particular. We used the same evaluation metrics as for the comparative methods.

Table 2: Age prediction results on the validation set from the 10-fold cross-validation, for the PREVENT-AD and OASIS3 datasets. Asterisk '*' indicates that the proposed method significantly outperformed the respective methods when compared using paired t-test (with $\alpha = 0.05$). Bold and blue indicate the **best** and second-best, respectively, in each column. LR = Linear Regression (Su et al., 2012); Support Vector Regression (SVR) (Awad & Khanna, 2015); DT = Decision Tree (De Ville, 2013); RT = Regression Tree (Loh, 2011); ET = Ensemble Tree (Che et al., 2011); NN = Neural Network (S.-C. Wang, 2003); MLP (Goodfellow et al., 2016); GCN (Kipf & Welling, 2016); GIN (K. Xu et al., 2018); GraphConv (Fey & Lenssen, 2019); ResGCN (Pei et al., 2022).

	PREVENT-AD				OASIS3			
Method	RMSE	MAE	PC	SC	RMSE	MAE	PC	SC
Naive model	4.63 ±1.08*	3.48 ±0.85	0	0	9.06 ±0.66*	7.19 ±0.56*	0	0
LR	5.12 ±0.48*	3.84 ±0.25*	**0.53 ±0.10**	**0.48 ±0.06**	7.76 ±0.53 *	6.15 ±0.44*	0.62 ±0.06*	0.62 ±0.05 *
SVR	4.92 ±0.39*	3.63 ±0.27	0.42 ±0.15	0.38 ±0.12	6.83 ±0.40	5.29 ±0.33	0.68 ±0.06	0.68 ±0.06*
DT	5.76 ±0.71*	4.23 ±0.57*	0.34 ±0.13	0.32 ±0.15	9.22 ±0.65*	7.27 ±0.54*	0.45 ±0.09*	0.44 ±0.09*
RT	5.78 ±0.79*	4.30 ±0.53*	0.35 ±0.11	0.31 ±0.14	9.18 ±0.62*	7.18 ±0.44*	0.45 ±0.09*	0.43 ±0.09*
ET	4.31 ±0.22	3.30 ±0.19	0.53 ±0.11	0.46 ±0.08	6.87 ±0.36	5.39 ±0.36	0.63 ±0.06	0.62 ±0.07
NN	5.97 ±0.64*	4.56 ±0.37*	0.09 ±0.12*	0.10 ±0.13*	8.54 ±1.28 *	6.13 ±0.40*	0.55 ±0.11	0.62 ±0.07
MLP	4.65 ±0.87*	3.6 ±0.71*	0.24 ±0.17*	0.24 ±0.17*	8.5 ±0.78*	6.14 ±0.44*	0.31 ±0.09*	0.41 ±0.07*
GCN	4.17 ±0.71	3.3 ±0.58	0.40 ±0.22	0.36 ±0.21	6.8 ±0.62	5.29 ±0.37	0.57 ±0.07*	0.60 ±0.06
GIN	4.57 ±0.74	3.64 ±0.66	0.39 ±0.26	0.34 ±0.22	9.17 ±3.27	6.12 ±0.80*	0.46 ±0.16*	0.57 ± 0.07
GraphConv	5.11 ±0.97*	3.97 ±0.81*	0.32 ±0.21*	0.29 ±0.17	8.17 ±0.52*	6.32 ±0.41*	0.28 ±0.08*	0.30 ±0.07*
ResGCN	4.31 ±0.58	3.43 ±0.48	0.41 ±0.22	0.37 ±0.19	7.64 ±1.07	5.39 ±0.50	0.53 ±0.11*	0.61 ±0.09
Proposed	**4.10 ±0.65**	**3.28 ±0.55**	0.45 ±0.20	0.38 ±0.21	**6.41 ±0.64**	**4.99 ±0.56**	**0.71 ±0.06**	**0.72 ±0.07**

4.2.5 Qualitative Results

We illustrate the scatter plots of the predictions vs. ground-truth data for the three tasks for the test set. We further report the significance of our results using a paired t-test between the folds.

5 Results and Discussion

5.1 Cross-validation and Unseen Test Set Experiments

Table 2 shows the results on the validation set of the 10-fold cross-validation for the age prediction task on the PREVENT-AD and OASIS3 datasets. The table is divided into three parts horizontally. The first part consists of the naive model. The mean RMSE is higher for the OASIS3 age prediction task due to its wider distribution compared to PREVENT-AD. However, the standard deviation of RMSE relative to its mean is smaller in OASIS3, possibly due to its larger dataset size. A larger dataset can provide more stable and reliable estimates of the true metric values, leading to reduced variation across cross-validation folds. The next set showcases the performance of 6 traditional ML methods. For the PREVENT-AD dataset, we observe that only the Ensemble Tree is able to outperform the naive baseline model. Where

Table 3: Cross-validation results by all methods for the MMSE prediction task on the validation set using the OASIS3 dataset. Bold and Blue indicate the **best** and second-best, respectively, in each column. For the fullforms and corresponding citation please refer to Table 2.

Method	RMSE	MAE	PC	SC
Naive model	2.38 ±0.52*	1.38 ±0.25	0	0
LR	2.94 ±0.31*	2.25 ±0.20*	0.24 ±0.11*	0.18 ±0.10*
SVR	**2.13 ±0.25**	**1.28 ±0.09***	**0.43 ±0.08**	**0.37 ±0.09**
DT	3.00 ±0.22 *	1.87 ±0.18 *	0.21 ±0.15 *	0.13 ±0.10 *
RT	3.09 ±0.41*	1.90 ±0.20*	0.18 ±0.13*	0.12 ±0.09*
ET	2.20 ±0.23	1.48 ±0.12	0.36 ±0.14	0.27 ±0.10
NN	3.31 ±0.52*	2.45 ±0.32*	0.23 ±0.32*	0.21 ±0.12*
MLP	2.76 ±0.44 *	1.71 ±0.17*	0.23 ±0.13*	0.30 ±0.08
GCN	2.25 ±0.48*	1.51 ±0.17*	0.36 ±0.08*	0.32 ± 0.08
GIN	2.94 ±0.56*	1.86 ±0.20*	0.28 ±0.13*	0.35 ±0.06
GraphConv	2.4 ±0.50*	1.63 ±0.20*	0.33 ±0.09*	0.31 ±0.08
ResGCN	2.34 ±0.39*	1.59 ±0.15*	0.35 ±0.08*	0.30 ±0.08
Proposed	2.20 ±0.43	1.44 ±0.22	0.42 ±0.08	0.31 ±0.09

Table 4: Performance evaluation for all the methods on 10% unseen test data. Bold and Blue indicate the **best** and second-best, respectively, in each column. For the fullforms and corresponding citation please refer to Table 2.

	PREVENT-AD				OASIS3							
	Age				Age				MMSE			
Method	RMSE	MAE	PC	SC	RMSE	MAE	PC	SC	RMSE	MAE	PC	SC
Naive model	6.17	5.65	0	0	9.46	7.21	0	0	3.02	1.71	0	0
LR	7.26	5.45	0.36	0.40	8.45	6.57	0.65	0.66	3.54	2.64	0.21	0.20
SVR	6.25	5.19	0.41	0.42	7.56	5.55	0.67	0.67	3.00	**1.63**	0.35	**0.35**
DT	7.50	6.15	0.20	0.17	10.86	8.37	0.29	0.28	3.40	1.92	0.20	0.19
RT	7.11	5.94	0.22	0.27	10.21	7.90	0.32	0.29	3.33	1.86	0.19	0.20
ET	5.54	4.69	0.37	0.39	7.35	5.57	0.64	0.62	**2.76**	1.64	**0.41**	0.31
NN	5.39	4.60	0.40	0.43	7.47	5.91	0.67	0.67	3.48	2.56	0.25	0.22
MLP	5.80	4.81	0.43	0.52	10.12	7.14	0.26	0.35	3.31	2.02	0.17	0.15
GCN	5.34	4.46	0.56	0.57	6.72	5.41	0.66	0.68	3.08	1.79	0.16	0.23
GIN	5.37	4.46	0.47	0.48	9.51	6.36	0.45	0.60	2.96	2.11	0.27	0.24
GraphConv	5.49	4.45	0.48	0.51	8.92	7.02	0.25	0.24	3.48	2.02	0.15	0.25
ResGCN	5.28	**4.39**	0.52	0.55	8.50	5.50	0.54	0.68	3.07	1.94	0.22	0.23
Proposed	**5.27**	4.42	**0.62**	**0.66**	**6.04**	**4.70**	**0.75**	**0.77**	2.89	1.76	0.34	0.27

as for the OASIS3 dataset, most of the traditional ML methods perform better than the naive classifier. The third part of the table is (mostly graph-based) DL methods. Most of the methods in this category show better performance than the naive model across both datasets. MLP, being the only non-graph DL

Table 5: Ablations on the prediction tasks. Each column represents the respective block removed from the pipeline. Bold and Blue indicate the **best** and second-best, respectively, in each row. Mod denotes a module ablation. We define: Mod 1 — no ResGCN layers; Mod 2 — no FC1 layer; Mod 3 — no CAB; Mod 4 — no skip connection; Mod i — no E in CAB; Mod ii (left) — no $G\mathbf{mm}^T$ and Mod ii (right) no G in $G\mathbf{mm}^T$; Mod iii— no Sigmoid in CAB; Mod iv — no MLPs in CAB.

	Model-level					CAB layer				
Metric	Proposed	Mod 1	Mod 2	Mod 3	Mod 4	Mod i	Mod ii		Mod iii	Mod iv
	PREVENT-AD (age)									
RMSE	**5.27**	5.76	5.30	5.65	5.62	6.47	6.38	6.47	6.02	6.44
MAE	**4.42**	4.68	4.50	4.7	4.68	5.14	5.09	5.16	4.87	5.14
PC	**0.62**	0.55	0.54	0.54	0.52	0.50	0.51	0.50	0.53	0.50
SC	**0.66**	0.57	0.58	0.57	0.56	0.51	0.51	0.51	0.58	0.51
	OASIS3 (age)									
RMSE	6.04	6.04	**5.78**	6.44	6.72	7.04	7.01	7.02	7.01	6.90
MAE	4.70	4.63	**4.47**	5.29	5.10	5.21	5.06	5.09	5.20	5.08
PC	0.75	0.73	**0.76**	0.71	0.68	0.64	0.64	0.64	0.65	0.65
SC	0.77	0.75	**0.77**	0.70	0.70	0.67	0.67	0.68	0.67	0.68
	OASIS3 (MMSE)									
RMSE	2.89	2.85	2.82	2.84	2.94	2.84	2.87	2.85	2.80	**2.79**
MAE	1.76	**0.83**	1.75	1.91	1.65	1.73	1.68	1.70	1.77	1.79
PC	0.34	0.35	0.35	0.36	**0.39**	0.34	0.35	0.34	0.38	0.38
SC	0.27	0.27	0.25	0.27	**0.33**	0.28	0.27	0.25	0.31	0.26

method, performs better than almost all of the traditional ML methods. The MLP's layered structure allows it to hierarchically extract and combine features, making it more expressive and better suited for modeling non-linear aging trajectories and interdependent biomarkers. Most of the graph DL methods perform better than traditional ML methods and MLP. This improved performance can be attributed to the ability of GNNs to model relational structures and population-level dependencies that are otherwise ignored in traditional or naive models. GNN-based methods incorporate subject-level connectivity information and leverage inter-subject similarities through graph representations. These methods can effectively capture non-linear interactions and high-dimensional spatial patterns in the data—especially useful in aging neuroimaging datasets, as brain connectivity is affected in aging. Furthermore, the inductive bias introduced by graph convolutions allows the models to generalize better to unseen subjects, particularly when brain regions or connectivity features exhibit complex dependencies. This capability likely explains their superior correlation with actual age, even when RMSE values are only moderately improved compared to simpler models. The proposed model consistently demonstrated superior performance, validating the strength of the architecture in capturing graph-structured population-level information across heterogeneous datasets. The proposed model integrates the strengths of MLPs (non-linear

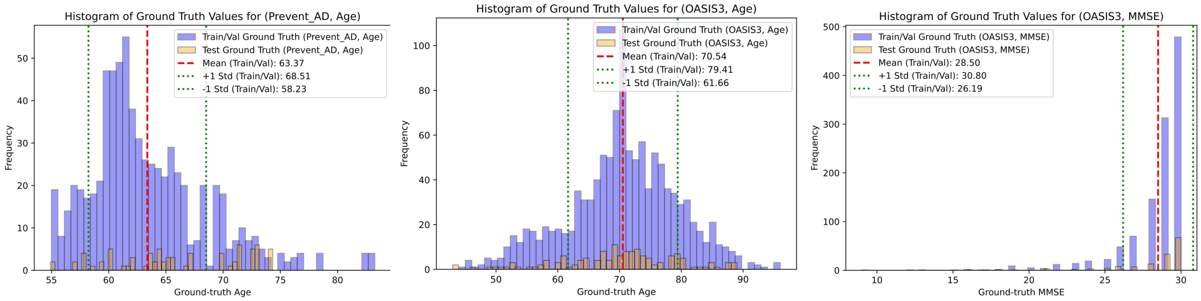


Figure 2: Histograms of the ground-truth values in the cross-validation (purple bars) and test (orange bars) sets for all datasets and tasks. The mean of each cross-validation distribution is marked by thick red dotted lines, while the standard deviation is indicated by thin green dotted lines. These visualizations help assess how well the training and test distributions align and provide insights into potential challenges in model training. For the PREVENT-AD dataset, the age distribution is notably skewed, indicating an imbalance in subject age representation. The OASIS3 dataset's MMSE distribution exhibits a power-law trend, where most samples cluster in the higher score range (i.e., cognitively normal individuals), with relatively few samples showing lower scores. The age distribution in the OASIS3 dataset is approximately symmetric and Gaussian, which is more favorable for training.

feature learning), ResGCNs (deep hierarchical graph representation), and a CAB (adaptive weighting of informative connections), enabling it to outperform all baseline methods by effectively capturing both individual and population-level patterns. The asterisk symbol shows where the proposed model performs significantly better based on the paired t-test $(\alpha = 0.05)$.

Similarly, Table 3 shows results on MMSE prediction on OASIS3 dataset. We can see a trend similar to Table 2, except that SVR performs best and the proposed method is mostly second best. Most of the traditional ML methods perform worse than the baseline, whereas the performance metrics of the DL methods are marginally better. This can be explained with a histogram of the ground-truth values, plotted for each dataset and task in the Figure. 2. As can be seen, MMSE has a much more skewed distribution than age does, potentially explaining the poor performance for the MMSE task. Although GNNs are effective for modeling complex relational patterns, MMSE scores are sparse, bounded, and heavily concentrated at higher values in healthy subjects, resulting in a skewed distribution and limited relational signal. For this type of endpoint, conventional models such as SVR and ensemble tree methods were able to capture the subtle variations adequately, and the added complexity of GNNs provided little additional benefit.

5.2 Results on 10% Unseen Test Data

We hypertune all the models during cross-validation, and then use the best setup for each model to evaluate its performance on the unseen test dataset. In these (single-fold) experiments, we train the model on all the cross-validation data and test it on the held-out test data. For each fold in the cross-validation, we stop the training at the epoch yielding the lowest loss on the validation set. For the held-out test data, we stop the training at the same epoch as that of the model corresponding to the lowest loss among the folds. For both datasets, as shown in Table 4, the proposed method outperformed all other models in age prediction in 7 out of 8 metrics, achieving the lowest prediction error (except it finished second in MAE for PREVENT-AD) while also demonstrating the highest correlation scores, indicating superior predictive accuracy. For MMSE prediction (in OASIS3), the proposed method still outperformed the other DL methods, but not the conventional ML methods of Ensemble Trees and (mostly) SVR. The narrow range of MMSE in the OASIS3 dataset (as reflected in the Naive model row and Figure 2) resulted in performance metric values that were close to each other among models.

We additionally performed these experiments on data from the second phase of the Alzheimer's Disease Neuroimaging Initiative, with a dataset size of 200 samples. While the proposed method still outperformed the rest of the methods, its prediction error was not substantially below the baseline, presumably due to the small dataset size (PC/SC were still 0.50/0.55 and 0.34/0.36 for age and MMSE, respectively).

5.3 Ablation Tests

Table 5 presents an ablation study on the prediction tasks on both datasets for the test set. For PREVENT-AD, when removing specific components such as the FC1 block, CAB block, skip connection, or GCN block, performance degrades across all datasets, confirming the effectiveness of the model's complete architecture. The trend remains generally consistent in OASIS3 (age); however, the original model without the FC1 component maintains the best scores, whereas other ablations cause notable declines in performance.

For OASIS3 (MMSE), the original model without the skip connection outperforms the other ablated versions in terms of correlations, though the performance drop is insubstantial. Removing any component results in only a marginal increase in RMSE and a decrease in correlation metrics, possibly due to the narrow range of MMSE values in OASIS3.

For age prediction, components that model global and relational structure are most critical. In particular, removing the GCN block (Mod 1) led to consistent degradation in correlation (e.g., PC drops from 0.62

to 0.55 in PREVENT-AD and from 0.75 to 0.73 in OASIS3), confirming that capturing inter-regional dependencies is essential for modeling brain aging. Similarly, the CAB module (Mod 3) showed substantial impact (e.g., RMSE increases from 6.04 to 6.44 in OASIS3), indicating that adaptive weighting of connections is important for identifying age-relevant connectivity patterns. This aligns with the biological premise that aging manifests as distributed, network-level changes. In contrast, for MMSE prediction, the contribution of individual modules is less pronounced. Ablation results show minimal variation across configurations (e.g., RMSE varies only from 2.79 to 2.94, and PC from 0.34 to 0.39). This suggests that MMSE prediction is less sensitive to architectural choices, likely due to the compressed and skewed distribution of MMSE scores (as shown in Figure 2), which limits the model's ability to leverage complex representations. Notably, removing the skip connection (Mod 4) slightly improves correlation (PC: 0.34 → 0.39), indicating that simpler representations may suffice for this task.

The ablation results generally emphasize that each architectural component plays a crucial role in maintaining predictive accuracy, with graph-based learning contributing significantly to overall model performance. The GCN block, in particular, appears essential for capturing complex dependencies, as its removal leads to a decline in the age prediction accuracy.

5.4 Qualitative Results

We show scatter plots of ground truth vs predictions on the held-out test set for all the datasets and tasks for qualitative evaluation of the performance. Figures 3, 4 and 5 show the scatter plots for age prediction in OASIS3 and PREVENT-AD and for MMSE prediction in OASIS3, respectively. The first two rows show the results by the DL methods, and the last two rows show those by traditional ML methods. The proposed method (top-left) shows the tightest clustering along the identity (45°) line, reflecting the highest correlation and lowest prediction error among all methods. These plots are consistent with the summary measures reported in the tables. A consistent trend among all methods is that the slope of the least-squares line is less than 45°, suggesting a compromise between the mean-predicting behavior of the naive model and ideal prediction.

5.5 Attention learned by CAB

We show the top five regions selected by the CAB in Table 6. For the age task for both datasets, the regions selected and attention values obtained are similar (but not identical) and are different from the MMSE task. As can be seen, the right hippocampus has the strongest weight for age in both datasets.

Table 6: Top regions identified across different tasks.

Id	PREVENT-AD Age	OASIS Age	OASIS MMSE
1	Right-Hippocampus	Right-Hippocampus	Left posterior cingulate cortex
2	Left-Pallidum	Left-Pallidum	Left banks of superior temporal sulcus
3	Right rostral anterior cingulate cortex	Right rostral anterior cingulate cortex	Right superior temporal cortex
4	Left postcentral cortex	Left postcentral cortex	Left pars triangularis cortex
5	Right cuneus cortex	Right pars triangularis	Right banks of superior temporal sulcus

Literature suggests that the hippocampus is indeed representative of aging (Aganj et al., 2023; Y. Xu et al., 2008); as we have seen in three datasets with aging populations, age was most strongly correlated with hippocampal connectivity (Aganj et al., 2023). Posterior cingulate cortex received the highest weighting for the MMSE task.

5.6 Hyperparameter and training time comparison

The output embedding sizes for the five modules were selected empirically through systematic hyperparameter sweeps during initial model development and benchmarking. The row 'Embedding size' in Table 7 shows the embedding sizes for the proposed and all the comparative methods.

As shown in Table 7 for the MMSE prediction task on OASIS3, most graph-based models maintain similar training times per epoch despite differences in parameter counts, while the proposed model shows a moderate increase in parameters with only a slight impact on runtime. In the Age prediction task on PREVENT-AD, the proposed model achieves a balanced trade-off, maintaining competitive training time with a moderate number of parameters compared to both lightweight (GraphConv) and heavier (GIN) architectures. For Age prediction on OASIS3, a larger variation is observed, where models such as GraphConv exhibit significantly higher parameter counts, whereas the proposed model remains relatively compact with competitive runtime. Overall, these results highlight that model complexity and computational cost are highly dataset- and task-dependent, and the proposed approach achieves a favorable balance between efficiency and capacity across different experimental settings. The proposed method is robust because it stays competitive in training time while using much fewer parameters in Age/OASIS3 than the other models. It is also consistently near the faster end of the table, which suggests a good balance between efficiency and model capacity.

Table 7: This table summarizes the selected hyperparameters and training efficiency across three setups. The learning rate was set to 0.0001 for all experiments, except for the age task on the OASIS3 dataset, where it was empirically set to 0.001. The proposed model is competitive in runtime, but its parameter count and speed vary by dataset, showing that performance depends strongly on the task.

	GCN	GIN	GraphConv	MLP	ResGCN	Proposed
MMSE / OASIS3						
Batch size	20	250	20	20	20	150
Embedding size	70	10	110	1	70	100
Total parameters	234,253	310,943	237,743	306,149	243,616	447,104
Time/epoch (sec)	0.39	0.39	0.39	0.31	0.56	0.41
Age / PREVENT-AD						
Batch size	20	120	60	60	40	120
Embedding size	30	50	10	10	10	1
Total parameters	96,433	316,627	28,383	175,913	32,456	109,590
Time/epoch (sec)	0.80	0.57	1.01	0.63	1.46	0.85
Age / OASIS3						
Batch size	20	250	20	20	20	90
Embedding size	50	10	110	1	70	1
Total parameters	238,418	104,347	1,061,783	151,300	451,196	23,430
Time/epoch (sec)	0.972	1.02	1.27	0.77	1.93	1.14

6 Conclusion

In this paper, we proposed a simple yet effective multi-branch model capable of capturing complementary information from structural brain connectivity, which we evaluated in the context of age and MMSE prediction. The configuration of input data and multiple operations from all four modules helped to learn better representations of each subject's graph. We have shown that our model often outperforms competing techniques on two publicly available datasets, while also ablating all the components of the model. Future work includes: the addition of interpretability to the models to find the brain subnetworks that are informative for the prediction task, as well as trying other (such as gated attention) graph convolution mechanisms; investigation of the suitability of as a subject-level attention score learned by the network, with the goal of quantifying how representative each subject is within the overall population;

systematic feature ablation studies across all methods to assess the contribution of each input modality to model performance; robustness analysis of our approach to domain shift and its generalizability by testing it on other clinical datasets after connectivity harmonization (Zhou et al., 2025).

7 Acknowledgments

Support for this research was provided by the National Institutes of Health (NIH), specifically the National Institute on Aging (RF1AG068261, R01AG068261).

Additional support was provided in part by the BRAIN Initiative Cell Census Network grant U01MH117023, the National Institute for Biomedical Imaging and Bioengineering (P41EB015896, R01EB023281, R01EB006758, R21EB018907, R01EB019956, P41EB030006), the National Institute on Aging (R56AG064027, R01AG064027, R01AG008122, R01AG016495, R01AG070988), the National Institute of Mental Health (R01MH121885, RF1MH123195), the National Institute for Neurological Disorders and Stroke (R01NS0525851, R21NS072652, R01NS070963, R01NS083534, U01NS086625, U24NS10059103, R01NS105820), the NIH Blueprint for Neuroscience Research (U01MH093765), part of the multi-institutional Human Connectome Project, and the Michael J. Fox Foundation for Parkinson's Research (MJFF-021226). Computational resources were provided through the Massachusetts Life Sciences Center.

B. Fischl is an advisor to DeepHealth, a company whose medical pursuits focus on medical imaging and measurement technologies. His interests were reviewed and are managed by Massachusetts General Hospital and Mass General Brigham in accordance with their conflict-of-interest policies.

References

Aganj, I., Lenglet, C., Jahanshad, N., Yacoub, E., Harel, N., Thompson, P. M., & Sapiro, G. (2011). A hough transform global probabilistic approach to multiple-subject diffusion mri tractography. *Medical Image Analysis*, *15*(4), 414–425.

Aganj, I., Lenglet, C., Sapiro, G., Yacoub, E., Ugurbil, K., & Harel, N. (2010). Reconstruction of the orientation distribution function in single-and multiple-shell q-ball imaging within constant solid angle. *Magnetic Resonance in Medicine*, *64*(2).

Aganj, I., Mora, J., Frau-Pascual, A., & Fischl, B. (2023). Exploratory correlation of the human structural connectome with non-mri variables in alzheimer's disease. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, *15*(4), e12511.

Aganj, I., Prasad, G., Srinivasan, P., Yendiki, A., Thompson, P. M., & Fischl, B. (2014). Structural brain network augmentation via kirchhoff's laws. *Joint Annual Meeting of ISMRM-ESMRMB, http://nmr.mgh.harvard.edu/~iman/ConductanceModel_ISMRM14_iman.pdf*, *22*, 2665.

Amoroso, N., La Rocca, M., Bruno, S., Maggipinto, T., Monaco, A., Bellotti, R., & Tangaro, S. (2017). Brain structural connectivity atrophy in alzheimer's disease. *arXiv preprint arXiv:1709.02369*.

Arrigo, A., Calamuneri, A., Milardi, D., Mormina, E., Gaeta, M., Corallo, F., Lo Buono, V., Chillemi, G., Marino, S., Cacciola, A., et al. (2019). Claustral structural connectivity and cognitive impairment in drug naıve parkinson's disease. *Brain Imaging and Behavior*, *13*, 933–944.

Arslan, S., Ktena, S. I., Glocker, B., & Rueckert, D. (2018). Graph saliency maps through spectral convolutional networks: Application to sex classification with brain connectivity. *Graphs in Biomedical Image Analysis and Integrating Medical Imaging and Non-Imaging Modalities: Second International Workshop, GRAIL 2018 and First International Workshop, Beyond MIC 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings 2*, 3–13.

Awad, M., & Khanna, R. (2015). Support vector regression. In *Efficient learning machines: Theories, concepts, and applications for engineers and system designers* (pp. 67–80). Springer.

Bahdanau, D. (2014). Neural machine translation by jointly learning to align and translate. *arXiv preprint arXiv:1409.0473*.

Bresson, X., & Laurent, T. (2017). Residual gated graph convnets. *arXiv preprint arXiv:1711.07553*.

Che, D., Liu, Q., Rasheed, K., & Tao, X. (2011). Decision tree and ensemble learning algorithms with their applications in bioinformatics. *Software tools and algorithms for biological systems*, 191–199.

Coelho, A., Fernandes, H. M., Magalhães, R., Moreira, P. S., Marques, P., Soares, J. M., Amorim, L., Portugal-Nunes, C., Castanho, T., Santos, N. C., et al. (2021). Reorganization of brain structural networks in aging: A longitudinal study. *Journal of Neuroscience Research*, *99*(5), 1354–1376.

Cole, J. H., Annus, T., Wilson, L. R., Remtulla, R., Hong, Y. T., Fryer, T. D., Acosta-Cabronero, J., Cardenas-Blanco, A., Smith, R., Menon, D. K., et al. (2022). Accurate age models for routine clinical mri examinations. *NeuroImage*, *249*, 118871.

Cole, J. H., Poudel, R. P., Tsagkrasoulis, D., Caan, M. W., Steves, C., Spector, T. D., & Montana, G. (2019). Age prediction using deep learning uncovers associated sequence variants. *Nature Communications*, *10*(1), 5409.

Damoiseaux, J. S. (2017). Effects of aging on functional and structural brain connectivity. *Neuroimage*, *160*, 32–40.

Dasoulas, G., Scaman, K., & Virmaux, A. (2021). Lipschitz normalization for self-attention layers with application to graph neural networks. *International Conference on Machine Learning*, 2456–2466.

De Ville, B. (2013). Decision trees. *Wiley Interdisciplinary Reviews: Computational Statistics*, *5*(6), 448–455.

Dennis, E. L., Jahanshad, N., McMahon, K. L., de Zubicaray, G. I., Martin, N. G., Hickie, I. B., Toga, A. W., Wright, M. J., & Thompson, P. M. (2013). Development of brain structural connectivity between ages 12 and 30: A 4-tesla diffusion imaging study in 439 adolescents and adults. *Neuroimage*, *64*, 671–684.

Dubost, F. (2020). Artificial intelligence with light supervision: Application to neuroimaging.

Faiyaz, A., Doyley, M. M., Schifitto, G., & Uddin, M. N. (2023). Artificial intelligence for diffusion mri-based tissue microstructure estimation in the human brain: An overview. *Frontiers in Neurology*, *14*, 1168833.

Feng, X., Yang, L., Zhang, Q., Cui, Z., Liu, S., Zhu, L., et al. (2024). Improving age prediction with anatomical feature attention-enhanced 3d-cnn. *Computers in Biology and Medicine*, *164*, 107380.

Feurer, M., Eggensperger, K., Falkner, S., Lindauer, M., & Hutter, F. (2020). Auto-sklearn 2.0: The next generation. *arXiv preprint arXiv:2007.04074*, *24*(8).

Fey, M., & Lenssen, J. E. (2019). Fast graph representation learning with pytorch geometric. *arXiv preprint arXiv:1903.02428*.

Fischl, B. (2012). Freesurfer. *Neuroimage*, *62*(2), 774–781.

Frau-Pascual, A., Augustinack, J., Varadarajan, D., Yendiki, A., Salat, D. H., Fischl, B., & Aganj, I. (2021). Conductance-based structural brain connectivity in aging and dementia. *Brain connectivity*, *11*(7), 566–583.

Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). *Deep learning* (Vol. 1). MIT press Cambridge.

Hagmann, P., Cammoun, L., Gigandet, X., Meuli, R., Honey, C. J., Wedeen, V. J., & Sporns, O. (2008). Mapping the structural core of human cerebral cortex. *PLoS biology*, *6*(7), e159.

He, Y., Chen, Z., & Evans, A. (2008). Structural insights into aberrant topological patterns of large-scale cortical networks in alzheimer's disease. *Journal of Neuroscience*, *28*(18), 4756–4766.

Hofman, M. A. (2015). *Evolution of the human brain: From matter to mind*. Springer.

Hu, J., Cao, L., Li, T., Dong, S., & Li, P. (2021). Gat-li: A graph attention network based learning and interpreting method for functional brain network classification. *BMC bioinformatics*, *22*, 1–20.

Jellinger, K. A. (2022). Morphological basis of parkinson disease-associated cognitive impairment: An update. *Journal of Neural Transmission*, *129*(8), 977–999.

Jiao, T., Li, F., Cui, Y., Wang, X., Li, B., Shi, F., Xia, Y., Zhou, Q., & Zeng, Q. (2023). Deep learning with an attention mechanism for differentiating the origin of brain metastasis using mr images. *Journal of Magnetic Resonance Imaging*, *58*(5), 1624–1635.

Joo, Y., Namgung, E., Jeong, H., Kang, I., Kim, J., Oh, S., Lyoo, I. K., Yoon, S., & Hwang, J. (2023). Brain age prediction using combined deep convolutional neural network and multi-layer perceptron algorithms. *Scientific Reports*, *13*(1), 22388.

Karlsgodt, K. H., Sun, D., & Cannon, T. D. (2010). Structural and functional brain abnormalities in schizophrenia. *Current directions in psychological science*, *19*(4).

Kazi, A., Farghadani, S., Aganj, I., & Navab, N. (2023). Ia-gcn: Interpretable attention based graph convolutional network for disease prediction. *International Workshop on Machine Learning in Medical Imaging*, 382–392.

Kazi, A., Markova, V., Kondamadugula, P. R., Liu, B., Adly, A., Faghihroohi, S., & Navab, N. (2022). Dg-gru: Dynamic graph based gated recurrent unit for age and gender prediction using brain imaging. *Medical Imaging 2022: Computer-Aided Diagnosis*, *12033*, 277–281.

Kazi, A., Mora, J., Fischl, B., Dalca, A. V., & Aganj, I. (2023). Multi-head graph convolutional network for structural connectome classification. *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 27–36.

Kazi, A., Shekarforoush, S., Arvind Krishna, S., Burwinkel, H., Vivar, G., Wiestler, B., Kortüm, K., Ahmadi, S.-A., Albarqouni, S., & Navab, N. (2019). Graph convolution based attention model for personalized disease prediction. *Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part IV 22*, 122–130.

Kazi, A., Shekarforoush, S., Kortuem, K., Albarqouni, S., Navab, N., et al. (2019). Self-attention equipped graph convolutions for disease prediction. *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, 1896–1899.

Kipf, T. N., & Welling, M. (2016). Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*.

Ktena, S. I., Parisot, S., Ferrante, E., Rajchl, M., Lee, M., Glocker, B., & Rueckert, D. (2018). Metric learning with spectral graph convolutions on brain connectivity networks. *NeuroImage*, *169*, 431–442.

Łabiak, A., Kazi, A., Pellegrini, C., Frau-Pascual, A., & Aganj, I. (2026). Multimodal connectivity-based cortical segmentation with graph neural networks. *Frontiers in Neuroscience*, *20*, 1729842.

LaMontagne, P. J., Benzinger, T. L., Morris, J. C., Keefe, S., Hornbeck, R., Xiong, C., Grant, E., Hassenstab, J., Moulder, K., Vlassenko, A. G., et al. (2019). Oasis-3: Longitudinal neuroimaging, clinical, and cognitive dataset for normal aging and alzheimer disease. *MedRxiv*, 2019–12.

Leoutsakos, J.-M., Gross, A., Jones, R., Albert, M., & Breitner, J. (2016). 'alzheimer's progression score': Development of a biomarker summary outcome for ad prevention trials. *The journal of prevention of Alzheimer's disease*, *3*(4), 229.

Lewis, J. D., O'Reilly, C., Bock, E., Theilmann, R. J., & Townsend, J. (2022). Aging-related differences in structural and functional interhemispheric connectivity. *Cerebral Cortex*, *32*(7), 1379–1389.

Li, Q., Han, Z., & Wu, X.-M. (2018). Deeper insights into graph convolutional networks for semi-supervised learning. *Proceedings of the AAAI conference on artificial intelligence*, *32*(1).

Li, X., Zhou, Y., Dvornek, N., Zhang, M., Gao, S., Zhuang, J., Scheinost, D., Staib, L. H., Ventola, P., & Duncan, J. S. (2021). Braingnn: Interpretable brain graph neural network for fmri analysis. *Medical Image Analysis*, *74*, 102233.

Li, X., Wang, Y., Wang, W., Huang, W., Chen, K., Xu, K., Zhang, J., Chen, Y., Li, H., Wei, D., et al. (2020). Age-related decline in the topological efficiency of the brain structural connectome and cognitive aging. *Cerebral cortex*, *30*(8), 4651–4661.

Liu, N., He, X., Laurent, T., Di Giovanni, F., Bronstein, M. M., & Bresson, X. (2024). A general graph spectral wavelet convolution via chebyshev order decomposition. *arXiv preprint arXiv:2405.13806*.

Liu, S., Park, J. H., & Yoo, S. (2020). Efficient and effective graph convolution networks. *Proceedings of the 2020 SIAM International Conference on Data Mining*, 388–396.

Loh, W.-Y. (2011). Classification and regression trees. *Wiley interdisciplinary reviews: data mining and knowledge discovery*, *1*(1), 14–23.

Meunier, D., Achard, S., Morcom, A., & Bullmore, E. (2009). Age-related changes in modular organization of human brain functional networks. *Neuroimage*, *44*(3), 715–723.

Moon, H. S., Mahzarnia, A., Stout, J., Anderson, R. J., Han, Z. Y., Tremblay, J. T., Badea, C. T., & Badea, A. (2024). Feature attention graph neural network for estimating brain age and identifying important neural connections in mouse models of genetic risk for alzheimer's disease. *Imaging Neuroscience*, *2*, 1–22.

Neudorf, J., Shen, K., & McIntosh, A. R. (2024). Reorganization of structural connectivity in the brain supports preservation of cognitive ability in healthy aging. *Network Neuroscience*, *8*(3), 837–859.

Oono, K., & Suzuki, T. (2019). Graph neural networks exponentially lose expressive power for node classification. *arXiv preprint arXiv:1905.10947*.

Pei, Y., Huang, T., Van Ipenburg, W., & Pechenizkiy, M. (2022). Resgcn: Attention-based deep residual modeling for anomaly detection on attributed networks. *Machine Learning*, *111*(2), 519–541.

Peng, H., Gong, W., Beckmann, C. F., Vedaldi, A., & Smith, S. M. (2021). Accurate age prediction with lightweight deep neural networks. *Medical Image Analysis*, *68*, 101871.

Peng, H., Gong, W., Beckmann, C. F., Vedaldi, A., & Smith, S. M. (2023). Age prediction using combined deep convolutional neural network and multi-layer perceptron algorithms. *Scientific Reports*, *13*(1), 20412.

Phan, A. V., Le Nguyen, M., Nguyen, Y. L. H., & Bui, L. T. (2018). Dgcnn: A convolutional neural network over large-scale labeled graphs. *Neural Networks*, *108*, 533–543.

Prescott, J. W., Doraiswamy, P. M., Gamberger, D., Benzinger, T., Petrella, J. R., & Network, D. I. A. (2022). Diffusion tensor mri structural connectivity and pet amyloid burden in preclinical autosomal dominant alzheimer disease: The dian cohort. *Radiology*, *302*(1), 143–150.

Qu, G., Zhou, Z., Calhoun, V. D., Zhang, A., & Wang, Y.-P. (2025). Integrated brain connectivity analysis with fmri, dti, and smri powered by interpretable graph neural networks. *Medical Image Analysis*, *103*, 103570.

Ranjbarzadeh, R., Bagherian Kasgari, A., Jafarzadeh Ghoushchi, S., Anari, S., Naseri, M., & Bendechache, M. (2021). Brain tumor segmentation based on deep learning and an attention mechanism using mri multi-modalities brain images. *Scientific Reports*, *11*(1), 1–17.

Reuter, M., Schmansky, N. J., Rosas, H. D., & Fischl, B. (2012). Within-subject template estimation for unbiased longitudinal image analysis. *Neuroimage*, *61*(4), 1402–1418.

Ryu, S., Kwon, O.-J., Park, J.-W., Bae, K.-H., Lee, S. H., Ryoo, H., & Yeom, J.-S. (2019). 3d-deep learning based automatic diagnosis of alzheimer's disease with joint mmse prediction using resting-state fmri. *Neuroinformatics*, *18*(1), 71–86.

Sebenius, I., Campbell, A., Morgan, S. E., Bullmore, E. T., & Liò, P. (2021). Multimodal graph coarsening for interpretable, mri-based brain graph neural network. *2021 IEEE 31st International Workshop on Machine Learning for Signal Processing (MLSP)*, 1–6.

Shi, F., Yap, P.-T., Gao, W., Lin, W., Gilmore, J. H., & Shen, D. (2012). Altered structural connectivity in neonates at genetic risk for schizophrenia: A combined study using morphological and white matter networks. *Neuroimage*, *62*(3), 1622–1633.

Sjöblom, L., Westerlund, D., Neimantaite, A. D., & Andersson, J. (2020). Braph for python: Graph theory and graph neural networks for brain connectivity analysis. *Emerging Topics in Artificial Intelligence 2020*, *11469*, 114691Y.

Sporns, O., Tononi, G., & Kötter, R. (2005). The human connectome: A structural description of the human brain. *PLoS Computational Biology*, *1*(4), e42.

Su, X., Yan, X., & Tsai, C.-L. (2012). Linear regression. *Wiley Interdisciplinary Reviews: Computational Statistics*, *4*(3), 275–294.

Tolan, E., & Isik, Z. (2018). Graph theory based classification of brain connectivity network for autism spectrum disorder. *Bioinformatics and Biomedical Engineering: 6th International Work-Conference, IWBBIO 2018, Granada, Spain, April 25–27, 2018, Proceedings, Part I 6*, 520–530.

Vaswani, A. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*.

Veličković, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., & Bengio, Y. (2017). Graph attention networks. *arXiv preprint arXiv:1710.10903*.

Wang, J., Khosrowabadi, R., Ng, K. K., Hong, Z., Chong, J. S. X., Wang, Y., Chen, C.-Y., Hilal, S., Venketasubramanian, N., Wong, T. Y., et al. (2018). Alterations in brain network topology and structural-functional connectome coupling relate to cognitive impairment. *Frontiers in Aging Neuroscience*, *10*, 404.

Wang, S., Tang, H., Himeno, R., Solé-Casals, J., Caiafa, C. F., Han, S., Aoki, S., & Sun, Z. (2024). Optimizing graph neural network architectures for schizophrenia spectrum disorder prediction using evolutionary algorithms. *Computer Methods and Programs in Biomedicine*, *257*, 108419.

Wang, S., Li, B. Z., Khabsa, M., Fang, H., & Ma, H. (2020). Linformer: Self-attention with linear complexity. *arXiv preprint arXiv:2006.04768*.

Wang, S.-C. (2003). Artificial neural network. In *Interdisciplinary computing in java programming* (pp. 81–100). Springer.

Wang, Y.-m., Zhang, Y.-j., Cai, X.-l., Yang, H.-x., Cheung, E. F., Chan, R. C., et al. (2020). Altered grey matter volume and white matter integrity in individuals with high schizo-obsessive traits, high schizotypal traits and obsessive-compulsive symptoms. *Asian Journal of Psychiatry*, *52*, 102096.

Wein, S., Malloni, W. M., Tomé, A. M., Frank, S. M., Henze, G.-I., Wüst, S., Greenlee, M. W., & Lang, E. W. (2021). A graph neural network framework for causal inference in brain networks. *Scientific reports*, *11*(1), 8061.

Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., & Philip, S. Y. (2020). A comprehensive survey on graph neural networks. *IEEE transactions on neural networks and learning systems*, *32*(1), 4–24.

Xie, T., & He, Y. (2012). Mapping the alzheimer's brain with connectomics. *Frontiers in psychiatry*, *2*, 77.

Xu, K., Hu, W., Leskovec, J., & Jegelka, S. (2018). How powerful are graph neural networks? *arXiv preprint arXiv:1810.00826*.

Xu, X., Wang, T., Li, W., Li, H., Xu, B., Zhang, M., Yue, L., Wang, P., & Xiao, S. (2021). Morphological, structural, and functional networks highlight the role of the cortical-subcortical circuit in individuals with subjective cognitive decline. *Frontiers in Aging Neuroscience*, *13*, 688113.

Xu, Y., Valentino, D. J., Scher, A. I., Dinov, I., White, L. R., Thompson, P. M., Launer, L. J., & Toga, A. W. (2008). Age effects on hippocampal structural changes in old men: The haas. *Neuroimage*, *40*(3), 1003–1015.

Zalesky, A., Cocchi, L., Fornito, A., Murray, M. M., & Bullmore, E. (2012). Connectivity differences in brain networks. *Neuroimage*, *60*(2), 1055–1062.

Zhang, F., Daducci, A., He, Y., Schiavi, S., Seguin, C., Smith, R. E., Yeh, C.-H., Zhao, T., & O'Donnell, L. J. (2022). Quantitative mapping of the brain's structural connectivity using diffusion mri tractography: A review. *Neuroimage*, *249*, 118870.

Zhang, S., Tong, H., Xu, J., & Maciejewski, R. (2019). Graph convolutional networks: A comprehensive review. *Computational Social Networks*, *6*(1), 1–23.

Zhang, X., He, L., Chen, K., Luo, Y., Zhou, J., & Wang, F. (2018). Multi-view graph convolutional network and its applications on neuroimage analysis for parkinson's disease. *AMIA annual symposium proceedings*, *2018*, 1147.

Zhang, Y., Zhan, L., Cai, W., Thompson, P., & Huang, H. (2019). Integrating heterogeneous brain networks for predicting brain disease conditions. *Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part IV 22*, 214–222.

Zhang, Y., He, X., Chan, Y. H., Teng, Q., & Rajapakse, J. C. (2023). Multi-modal graph neural network for early diagnosis of alzheimer's disease from smri and pet scans. *Computers in Biology and Medicine*, *164*, 107328.

Zhang, Y., Teng, Q., He, X., Niu, T., Zhang, L., Liu, Y., & Ren, C. (2023). Attention-based 3d cnn with multi-layer features for alzheimer's disease diagnosis using brain images. *2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, 1–4.

Zhao, X., Wu, P., Chen, L., Liu, J., Li, S., & Yao, D. (2024). Age prediction using interpretable multi-feature-based convolutional neural network in mild traumatic brain injury. *NeuroImage*, *295*, 120644.

Zheng, K., Yu, S., Chen, L., Dang, L., & Chen, B. (2024). Bpi-gnn: Interpretable brain network-based psychiatric diagnosis and subtyping. *NeuroImage*, *292*, 120594.

Zhou, Z., Fischl, B., & Aganj, I. (2025). Harmonization of structural brain connectivity through distribution matching [e70257 HBM-24-0961.R2]. *Human Brain Mapping*, *46*(9), e70257. https://doi.org/https://doi.org/10.1002/hbm.70257

Zhou, Z., Li, J., Williams, J., Fischl, B., & Aganj, I. (2026). Deep-learning cortical registration guided by structural and diffusion mri and connectivity. *International Symposium on Biomedical Imaging (ISBI)*.

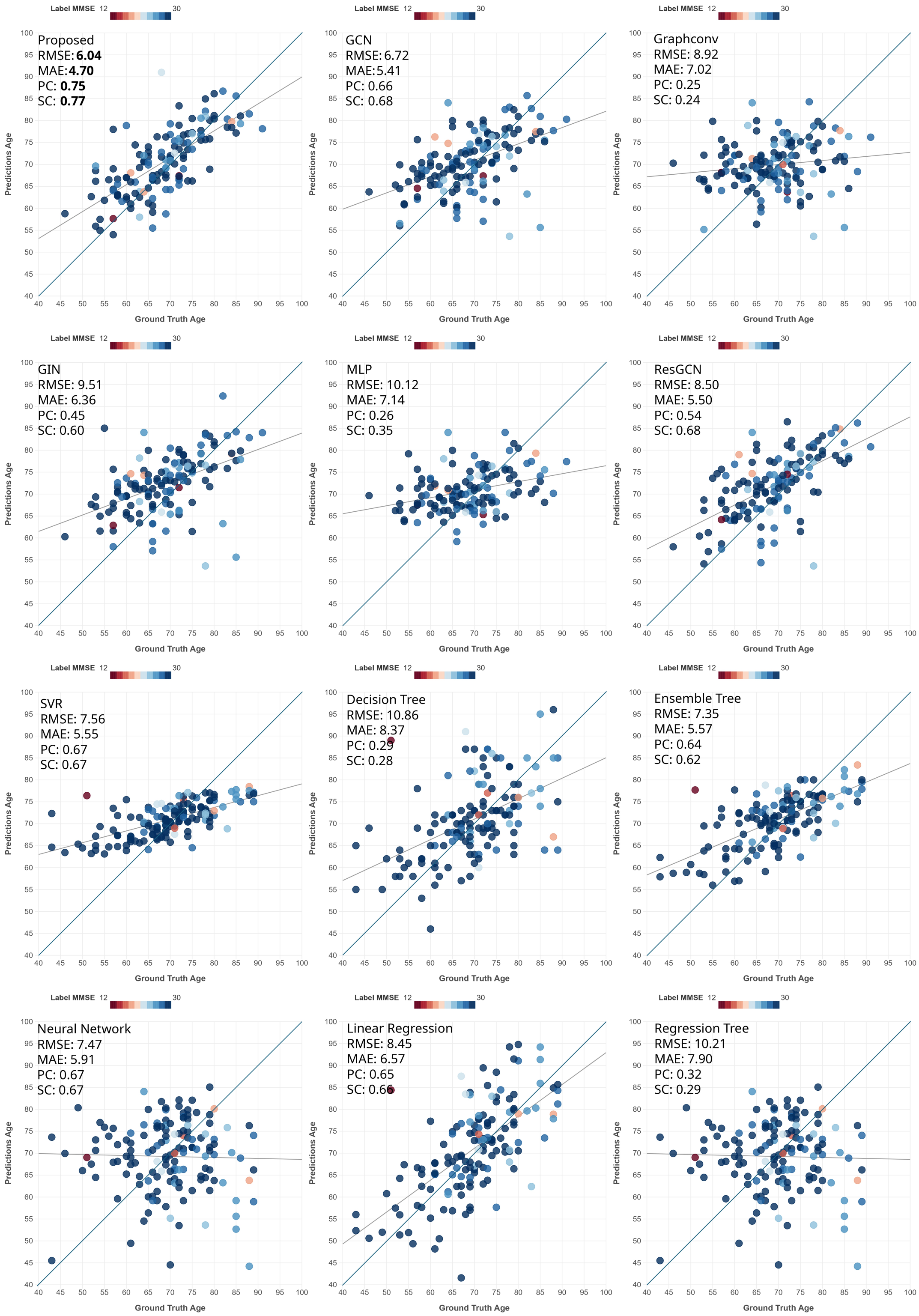


Figure 3: Scatter plots comparing predicted age to ground-truth age on the OASIS3 dataset. Each dot represents a subject in the test set. The darker line is the identity line, whereas the lighter line is the least-squares line. The color indicates the MMSE score.

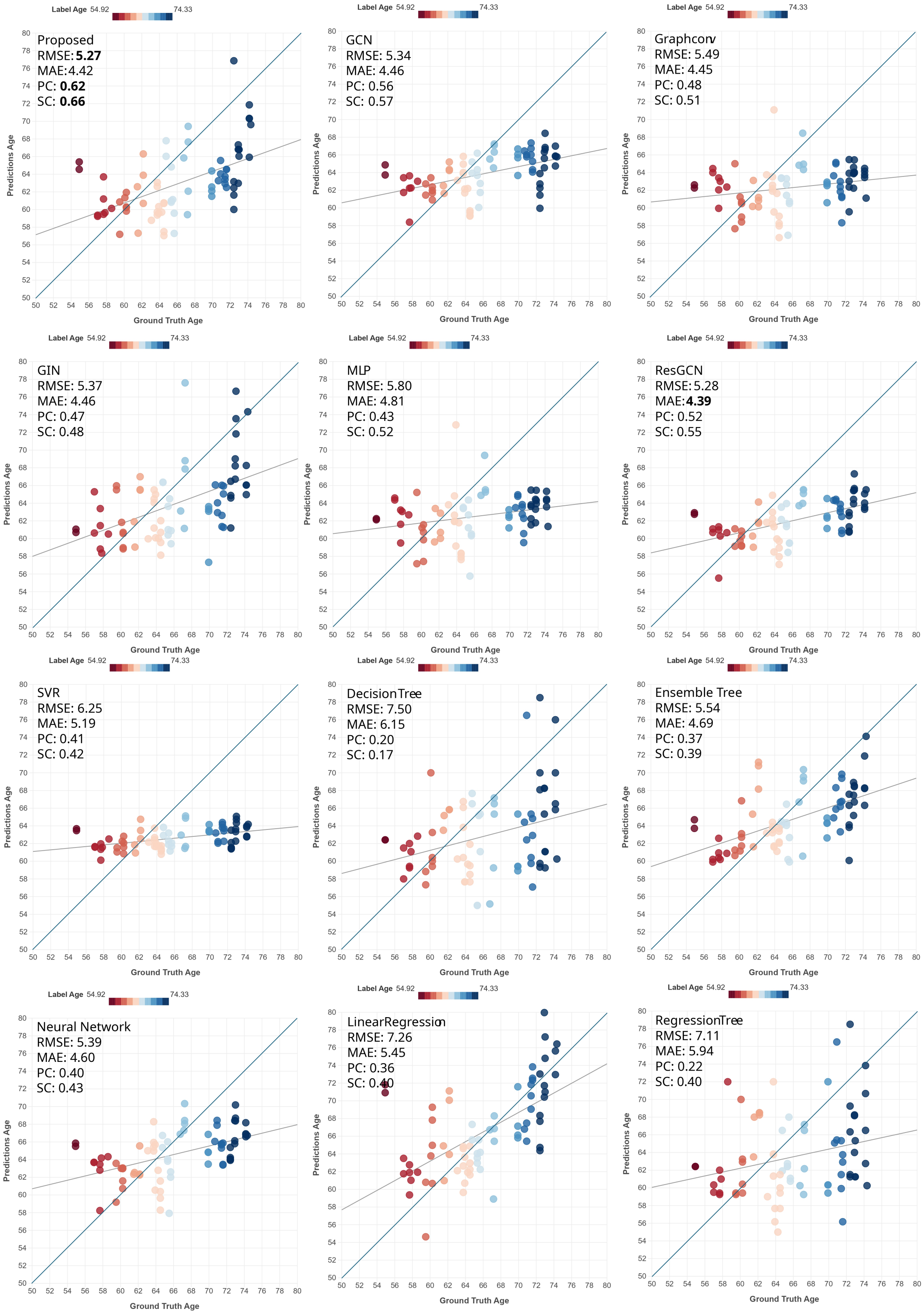


Figure 4: Scatter plots of prediction performance for the test set of the PREVENT-AD dataset for the task of age prediction. The darker line is the identity line, whereas the lighter line is the least-squares line. The color indicates the ground-truth age.

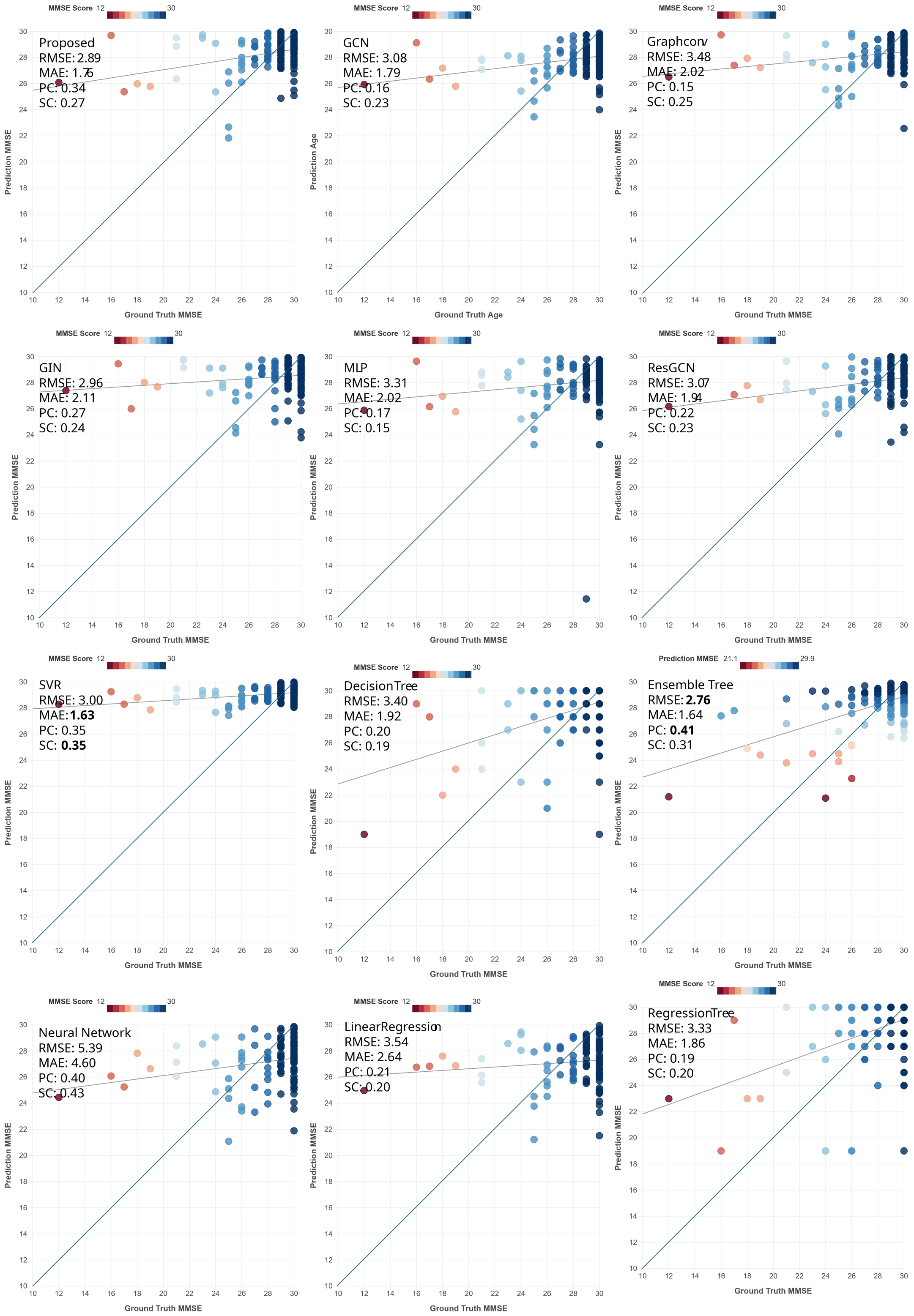
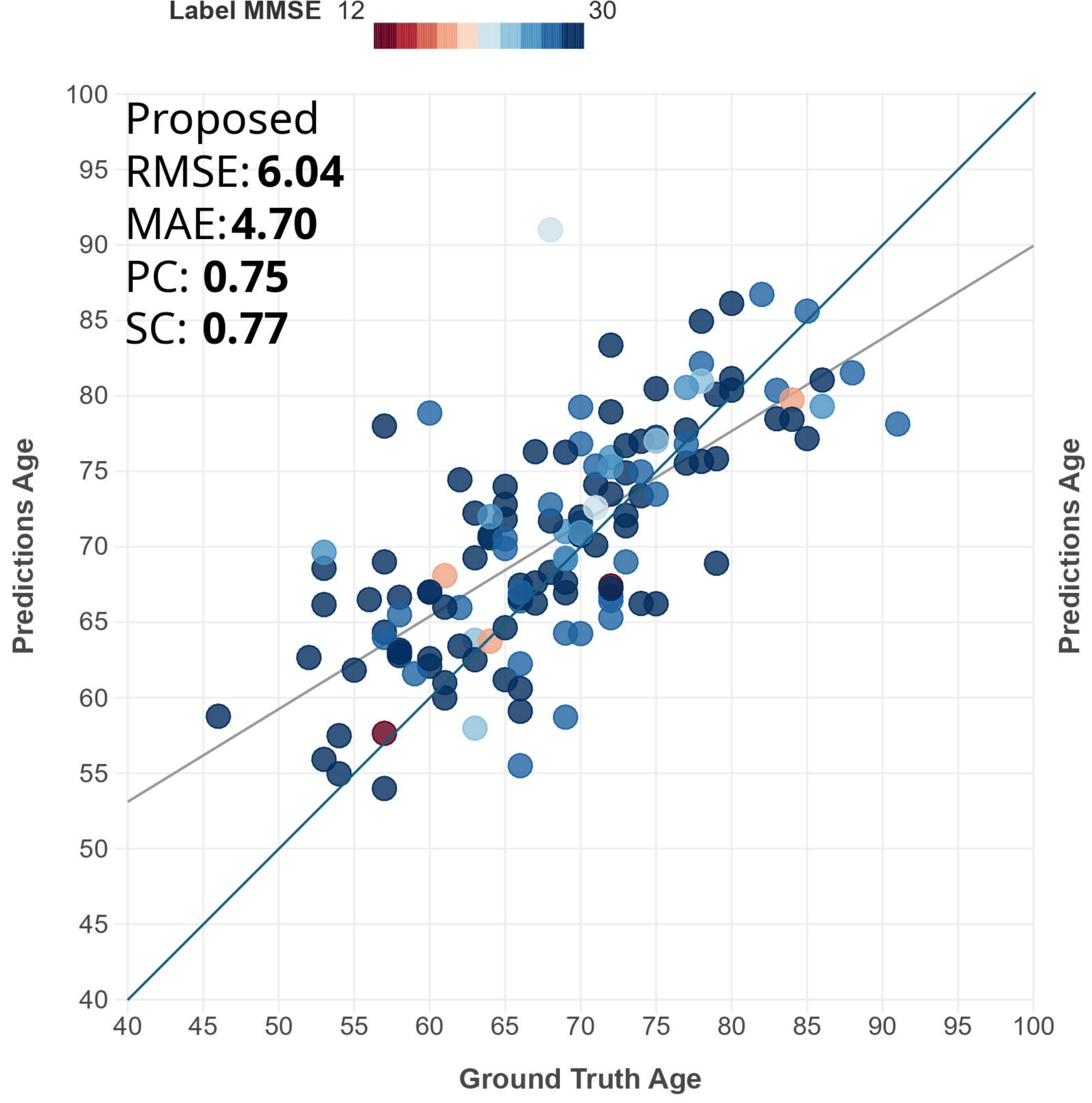
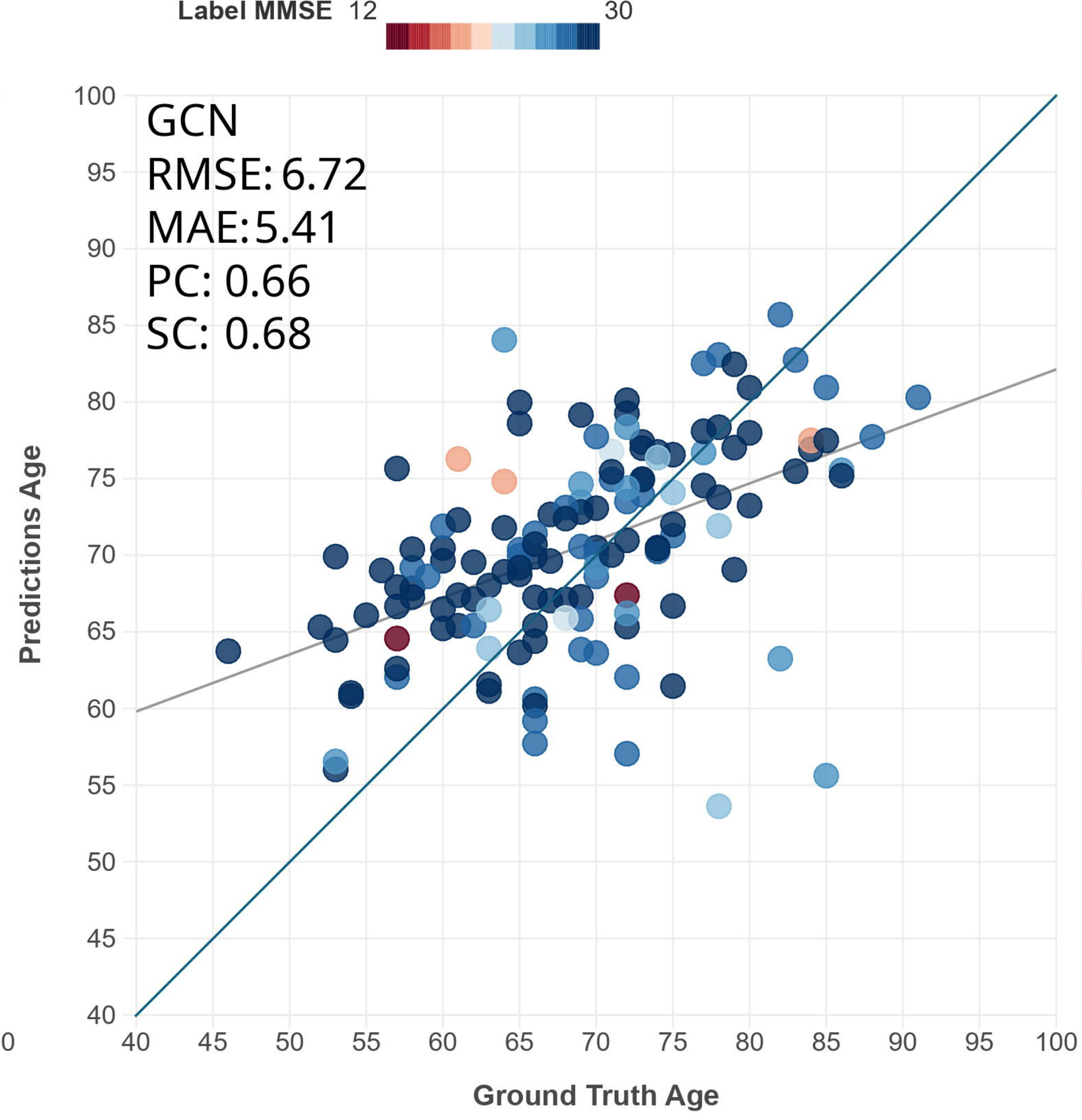


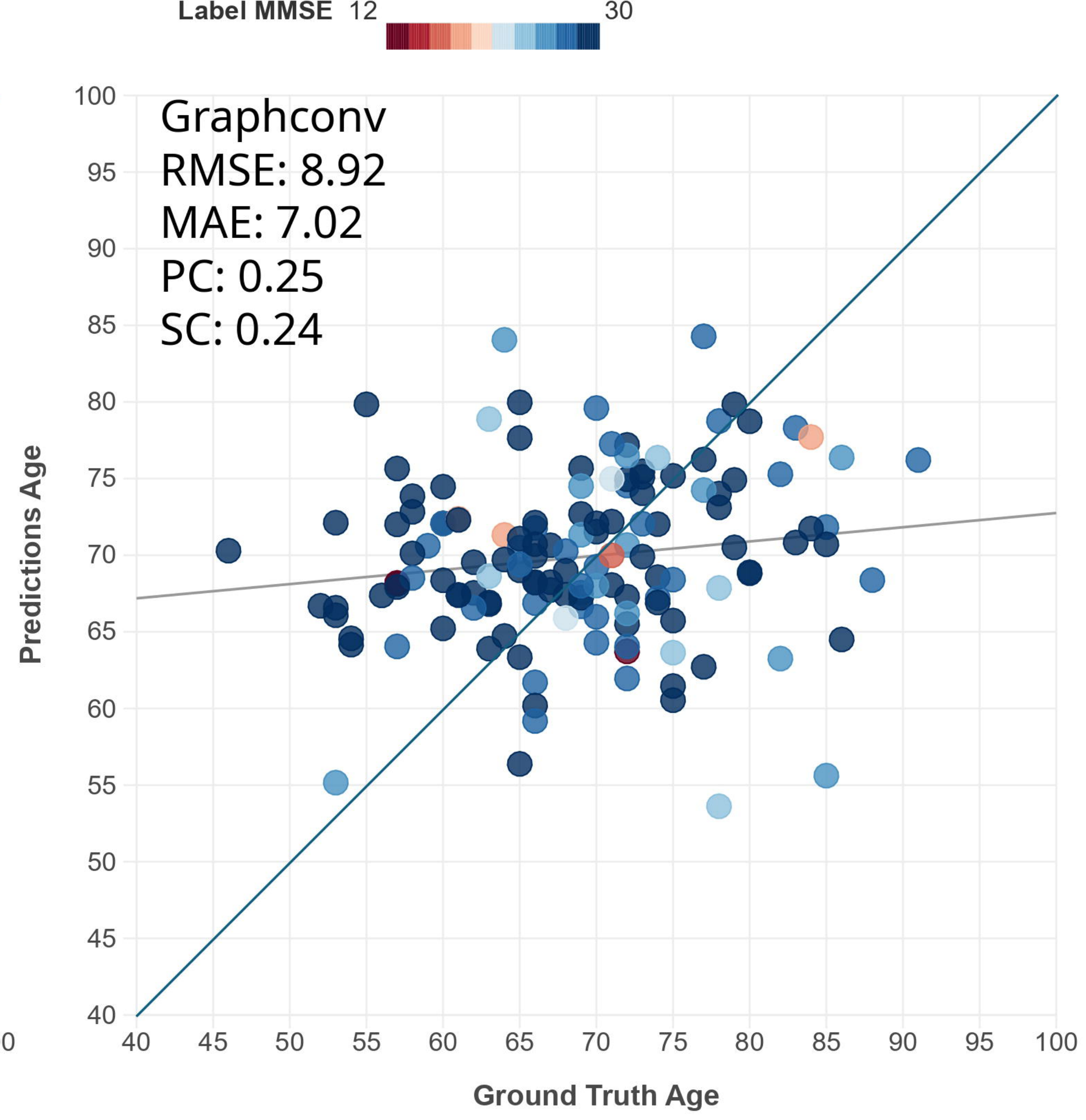
Figure 5: Scatter plots for test set on OASIS3 dataset for the task of MMSE prediction. The darker line is the identity line, whereas the lighter line is the least-squares line. The color indicates the ground-truth MMSE.

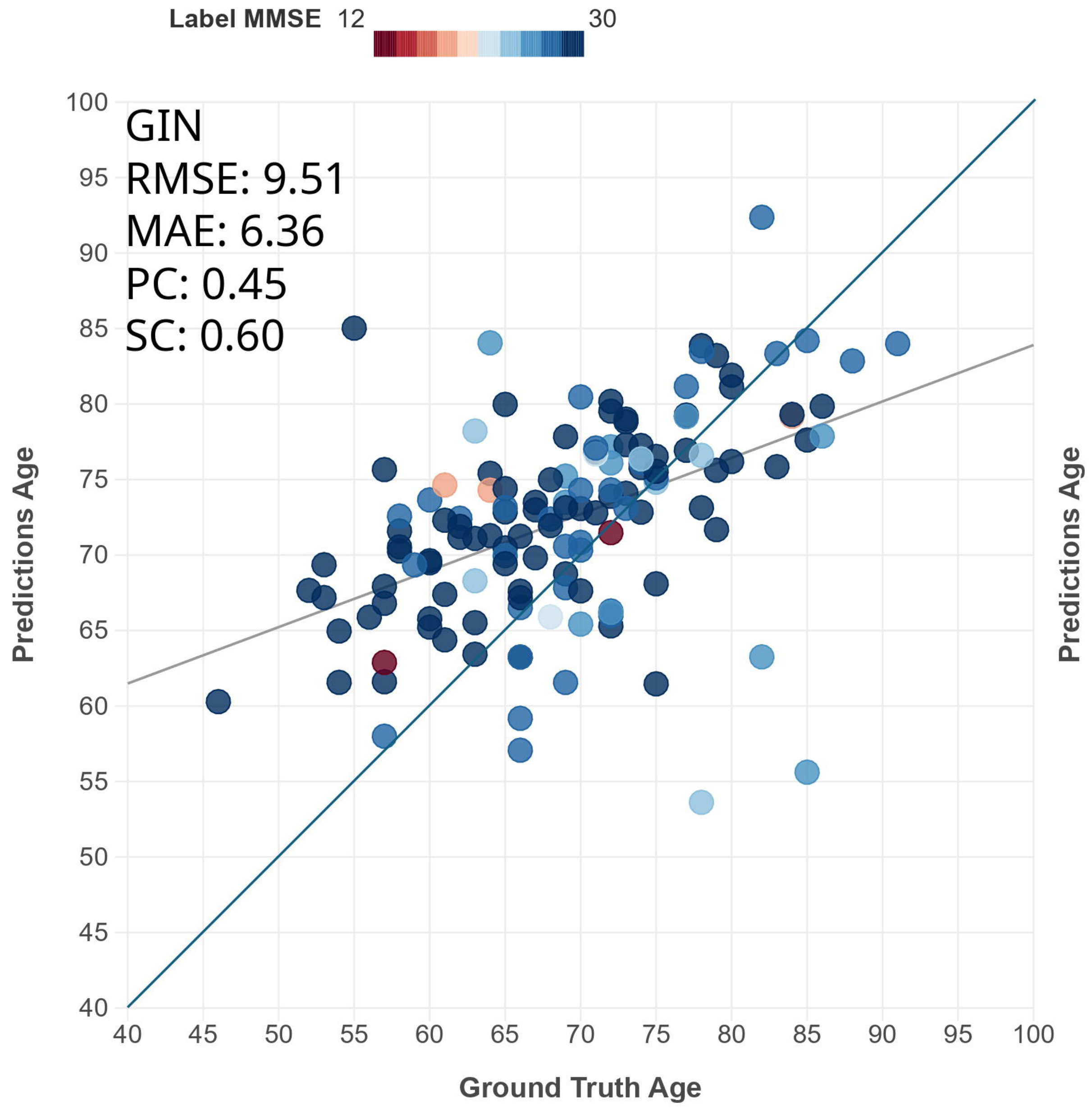
Pre-processing
Diffusion MRI + Tractography
FA, ADC, Node degree
Pre-processing
FreeSurfer Altas
1
2
85
1
85
X
G
End-to-end pipeline, f_θ
ResGCN
ResGCN
$\mathbf{h}^1$
$\mathbf{h}^2$
FC2
$\mathbf{h}^6$
Module 1
$vec(\mathbf{X})$
FC1
$\mathbf{h}^3$
Module 2
Connectivity Attention Block (CAB)
h^4
$\mathbf{h}^5$
Module 3
Module 4
FC5
$\hat{y}$
L
Connectivity Attention Block, $C_{\phi,\mathbf{m}}$
G
E
h^4
Module i
m
$\mathbf{m}^T$
$\mathbf{G}'$
X
Module ii
FC3
Module iii
FC4
Module iv
$\mathbf{h}^5$
Flattening
Concatenation
Matrix multiplication
mask
ReLU
Sigmoid

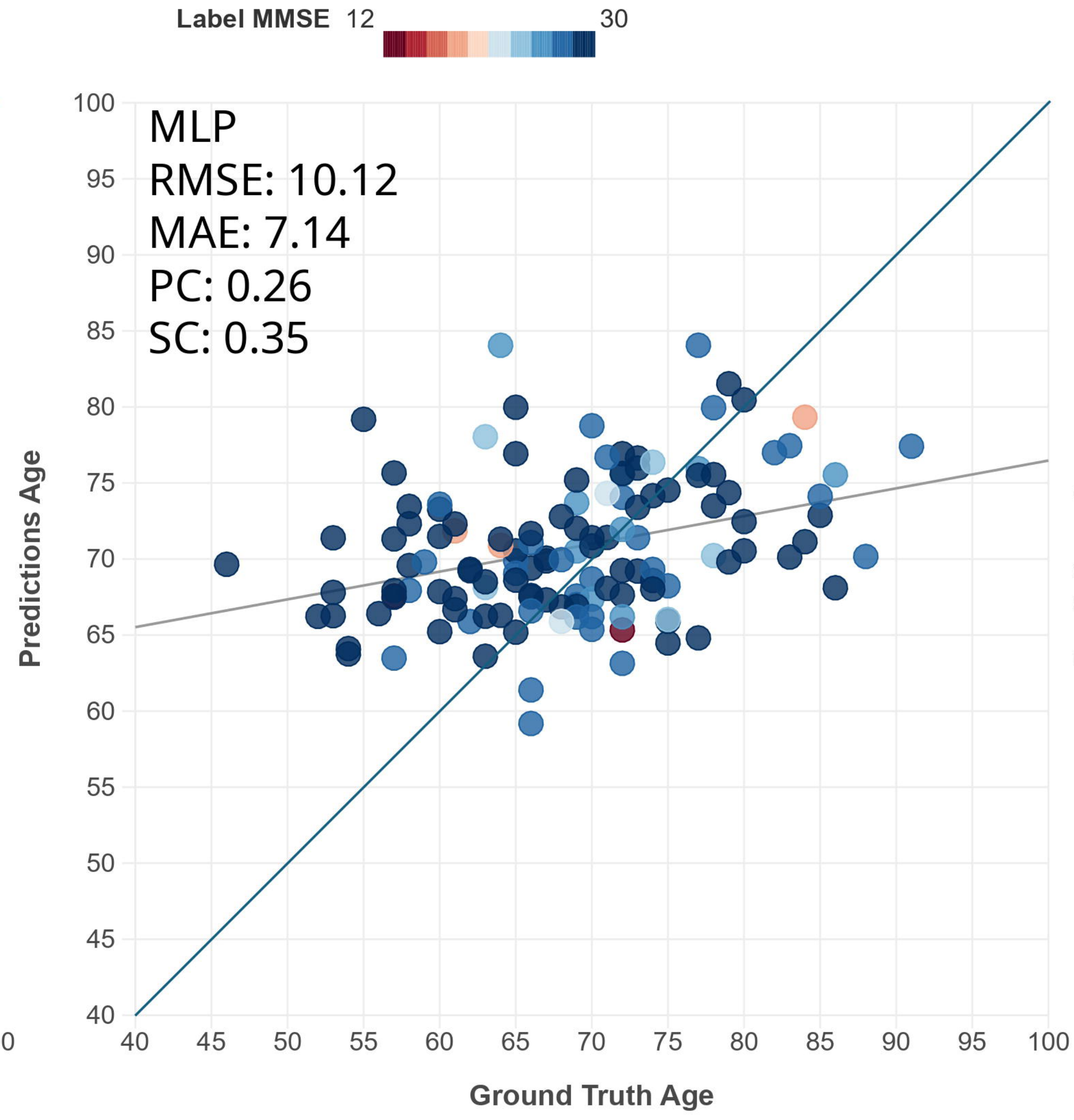
Histogram of Ground Truth Values for (Prevent_AD, Age)
Train/Val Ground Truth (Prevent_AD, Age)
Test Ground Truth (Prevent_AD, Age)
Mean (Train/Val): 63.37
+1 Std (Train/Val): 68.51
-1 Std (Train/Val): 58.23
Frequency
0
10
20
30
40
50
55
60
65
70
75
80
Ground-truth Age
Histogram of Ground Truth Values for (OASIS3, Age)
Train/Val Ground Truth (OASIS3, Age)
Test Ground Truth (OASIS3, Age)
Mean (Train/Val): 70.54
+1 Std (Train/Val): 79.41
-1 Std (Train/Val): 61.66
Frequency
0
20
40
60
80
100
50
60
70
80
90
Ground-truth Age
Histogram of Ground Truth Values for (OASIS3, MMSE)
Train/Val Ground Truth (OASIS3, MMSE)
Test Ground Truth (OASIS3, MMSE)
Mean (Train/Val): 28.50
+1 Std (Train/Val): 30.80
-1 Std (Train/Val): 26.19
Frequency
0
100
200
300
400
500
10
15
20
25
30
Ground-truth MMSE

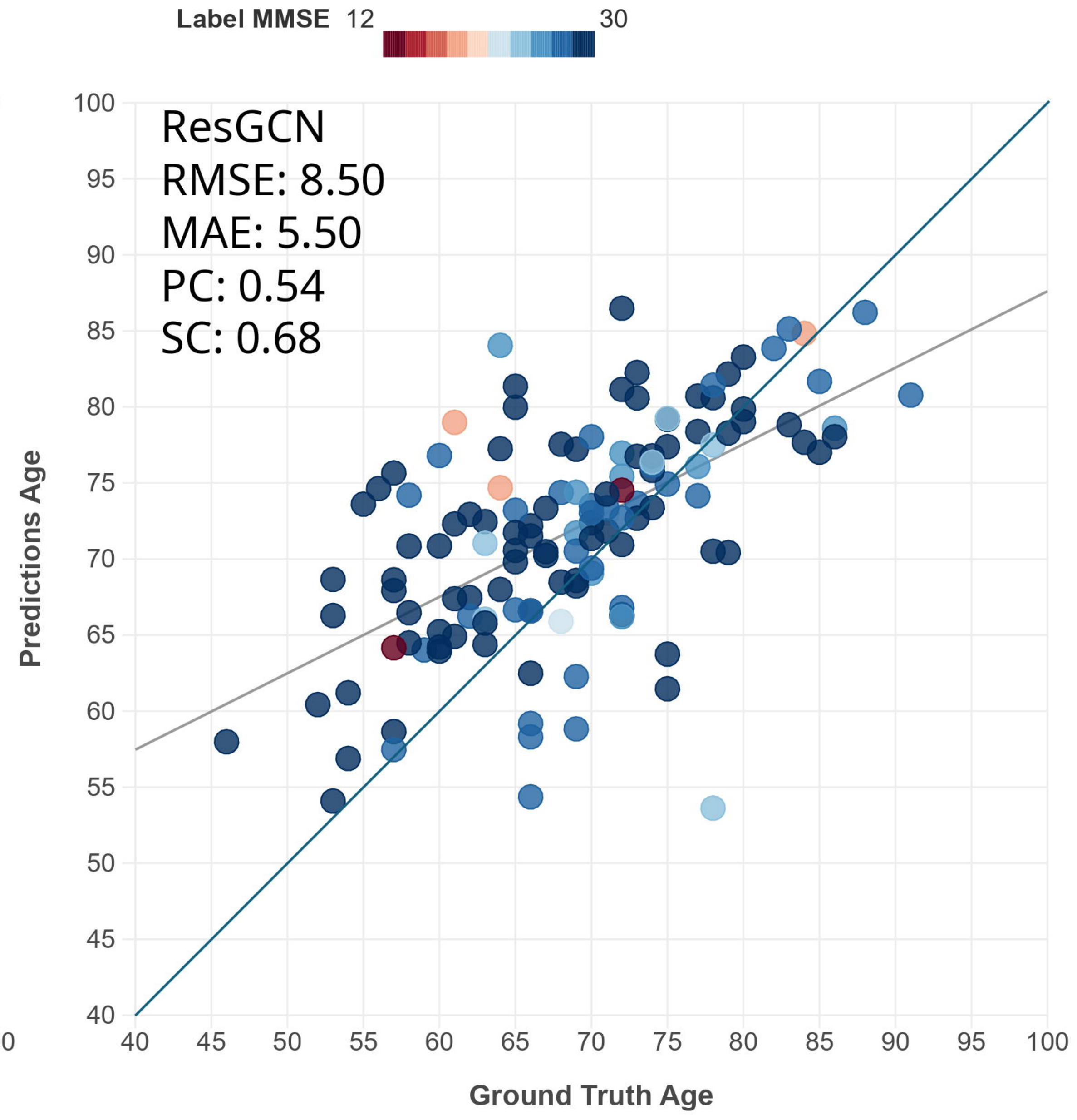

Label MMSE 12 30
Proposed
RMSE: **6.04**
MAE: **4.70**
PC: **0.75**
SC: **0.77**
Predictions Age
Ground Truth Age
40 45 50 55 60 65 70 75 80 85 90 95 100

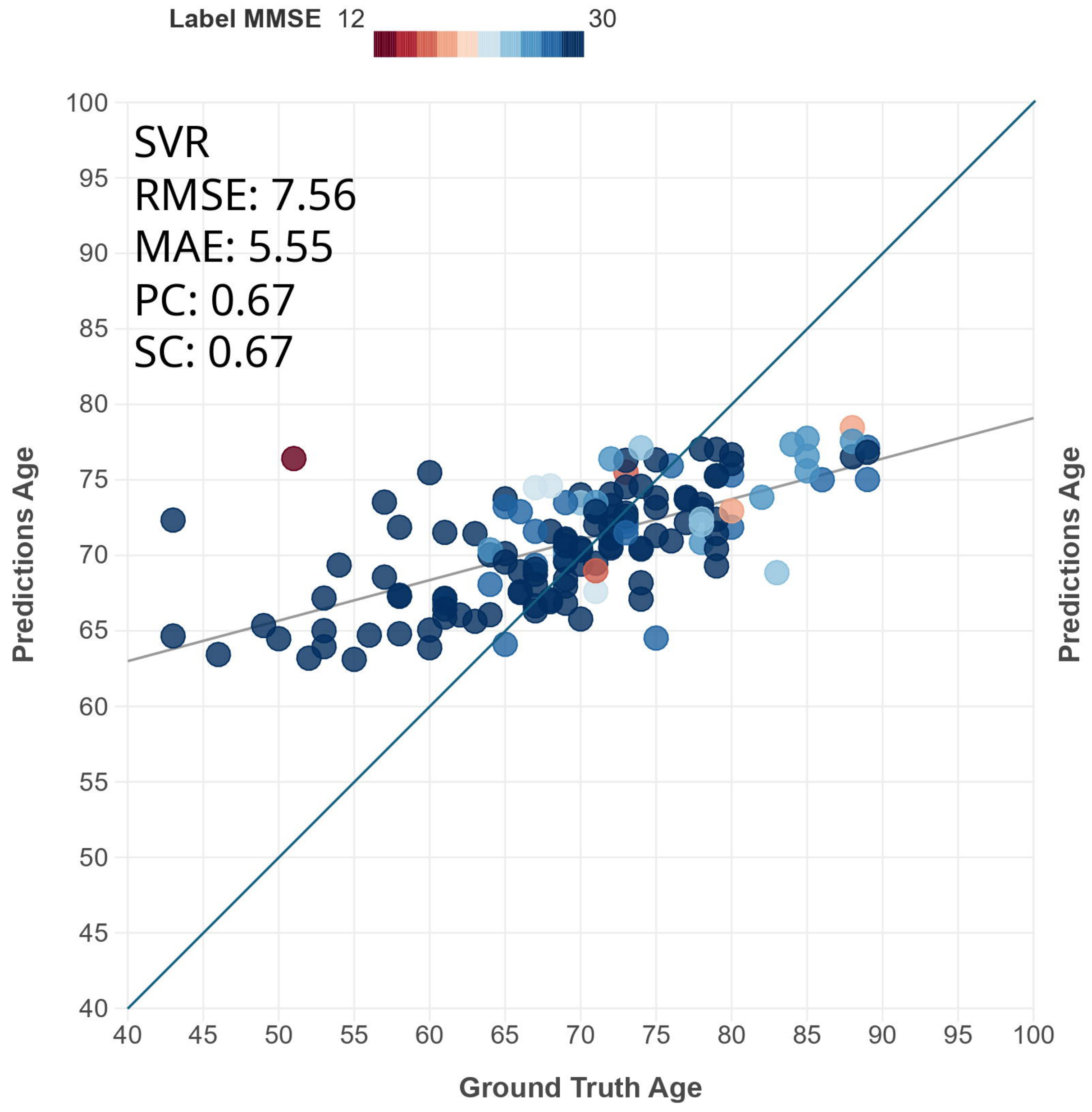

Label MMSE 12 30
GCN
RMSE: 6.72
MAE: 5.41
PC: 0.66
SC: 0.68
Predictions Age
Ground Truth Age
40 45 50 55 60 65 70 75 80 85 90 95 100

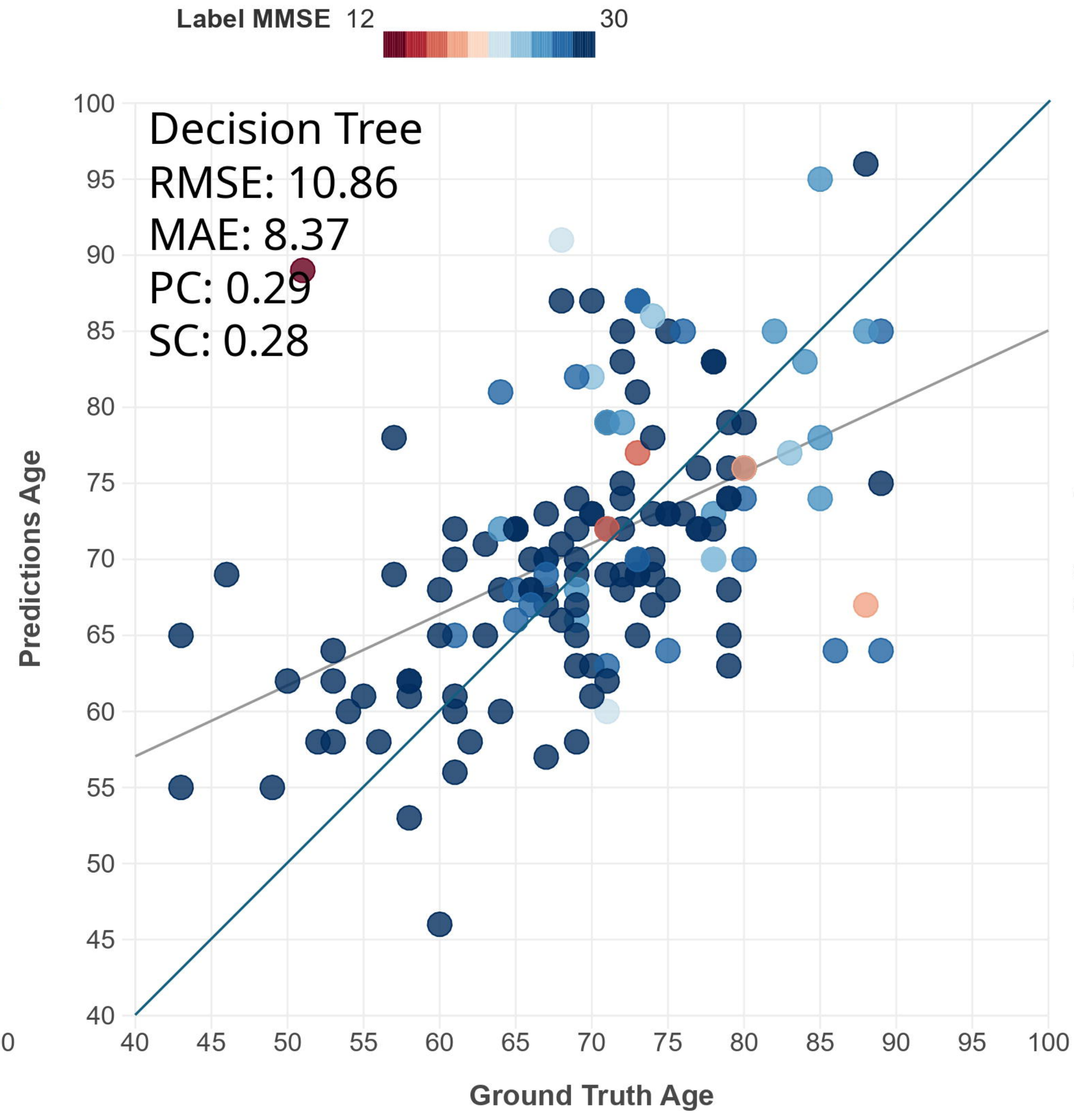

Label MMSE 12 30
Graphconv
RMSE: 8.92
MAE: 7.02
PC: 0.25
SC: 0.24
Predictions Age
Ground Truth Age
40 45 50 55 60 65 70 75 80 85 90 95 100

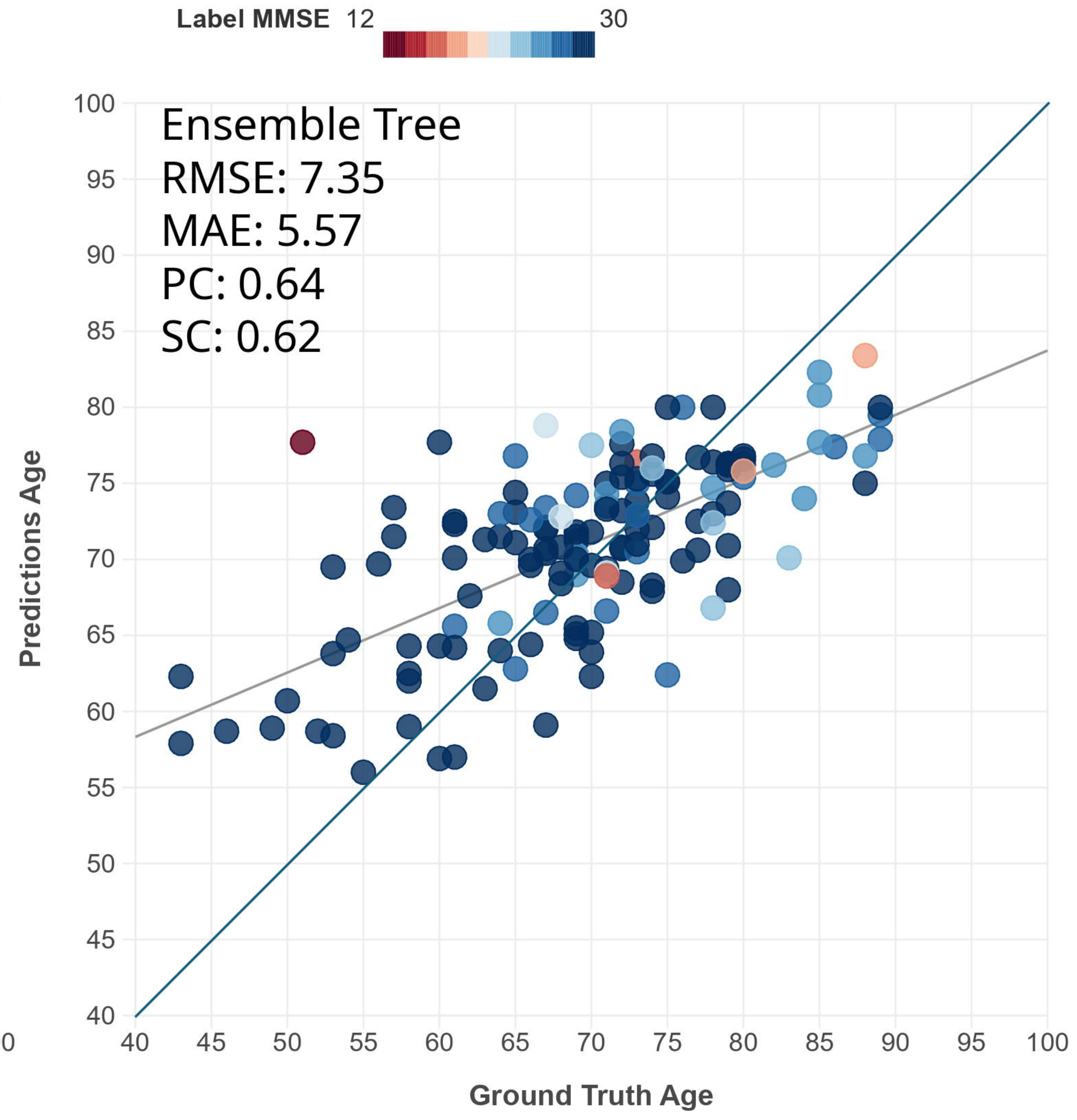

Label MMSE 12 30
GIN
RMSE: 9.51
MAE: 6.36
PC: 0.45
SC: 0.60
Predictions Age
Ground Truth Age
40 45 50 55 60 65 70 75 80 85 90 95 100

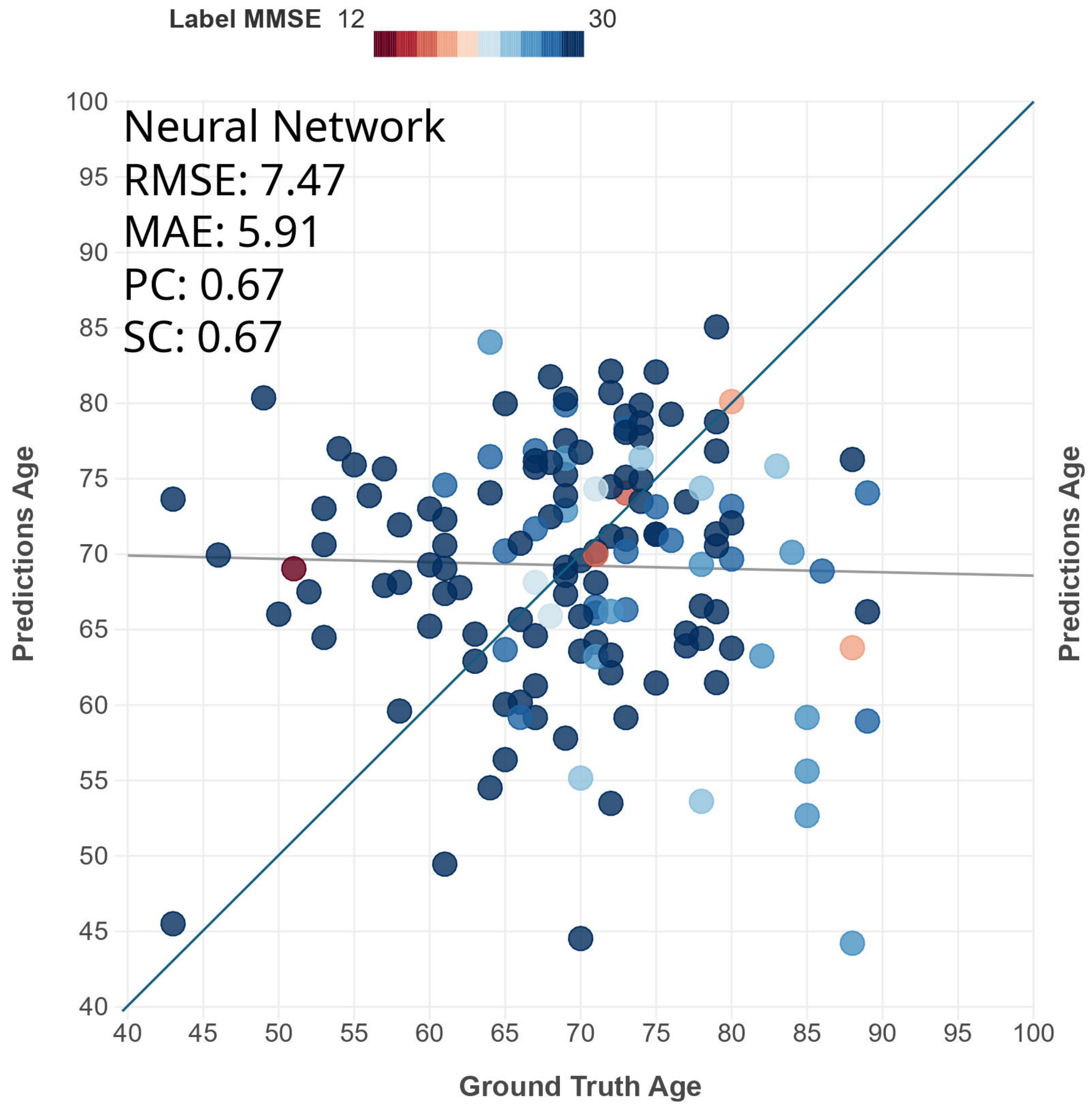

Label MMSE 12 30
MLP
RMSE: 10.12
MAE: 7.14
PC: 0.26
SC: 0.35
Predictions Age
Ground Truth Age
40 45 50 55 60 65 70 75 80 85 90 95 100

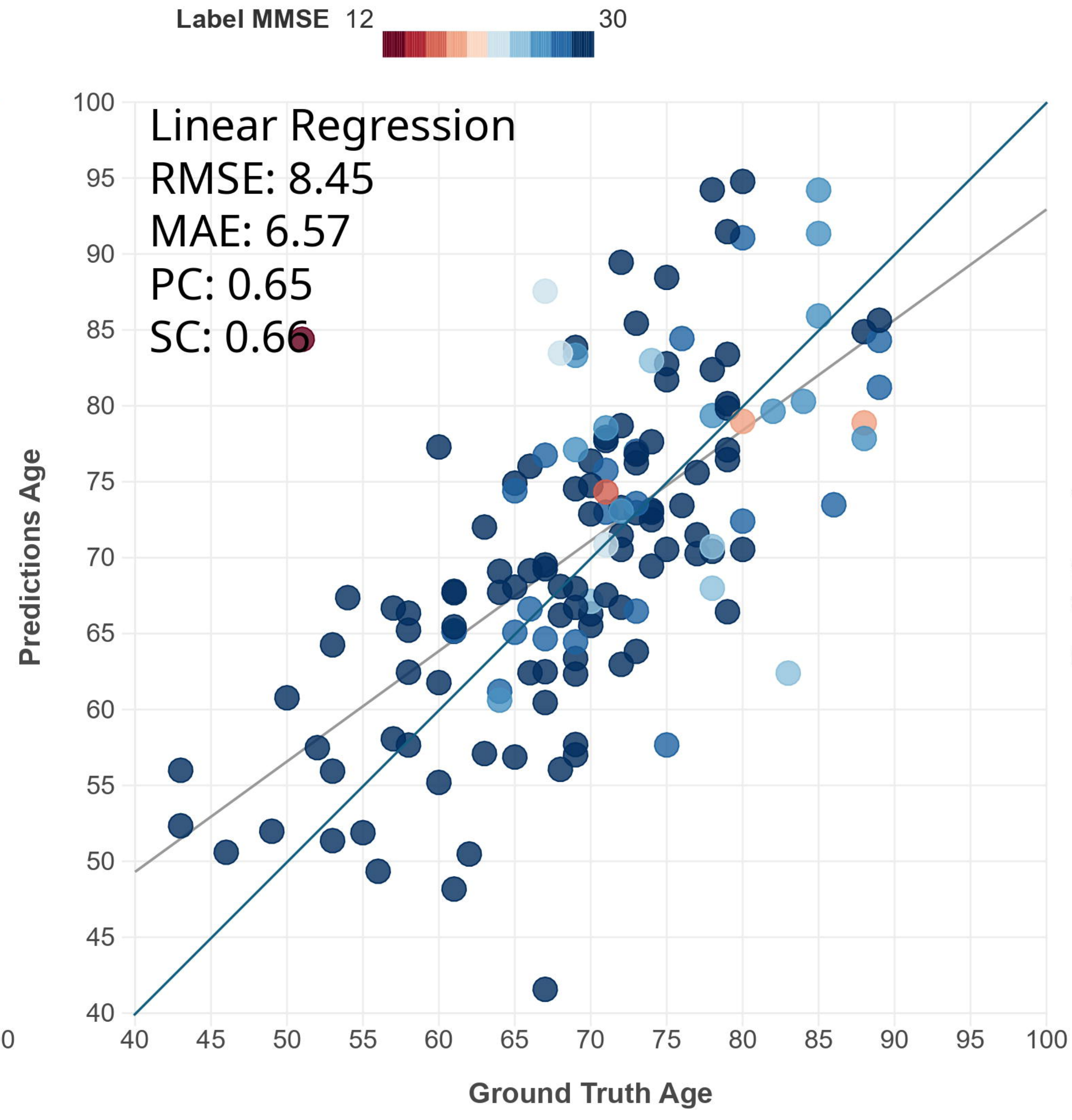

Label MMSE 12 30
ResGCN
RMSE: 8.50
MAE: 5.50
PC: 0.54
SC: 0.68
Predictions Age
Ground Truth Age
40 45 50 55 60 65 70 75 80 85 90 95 100

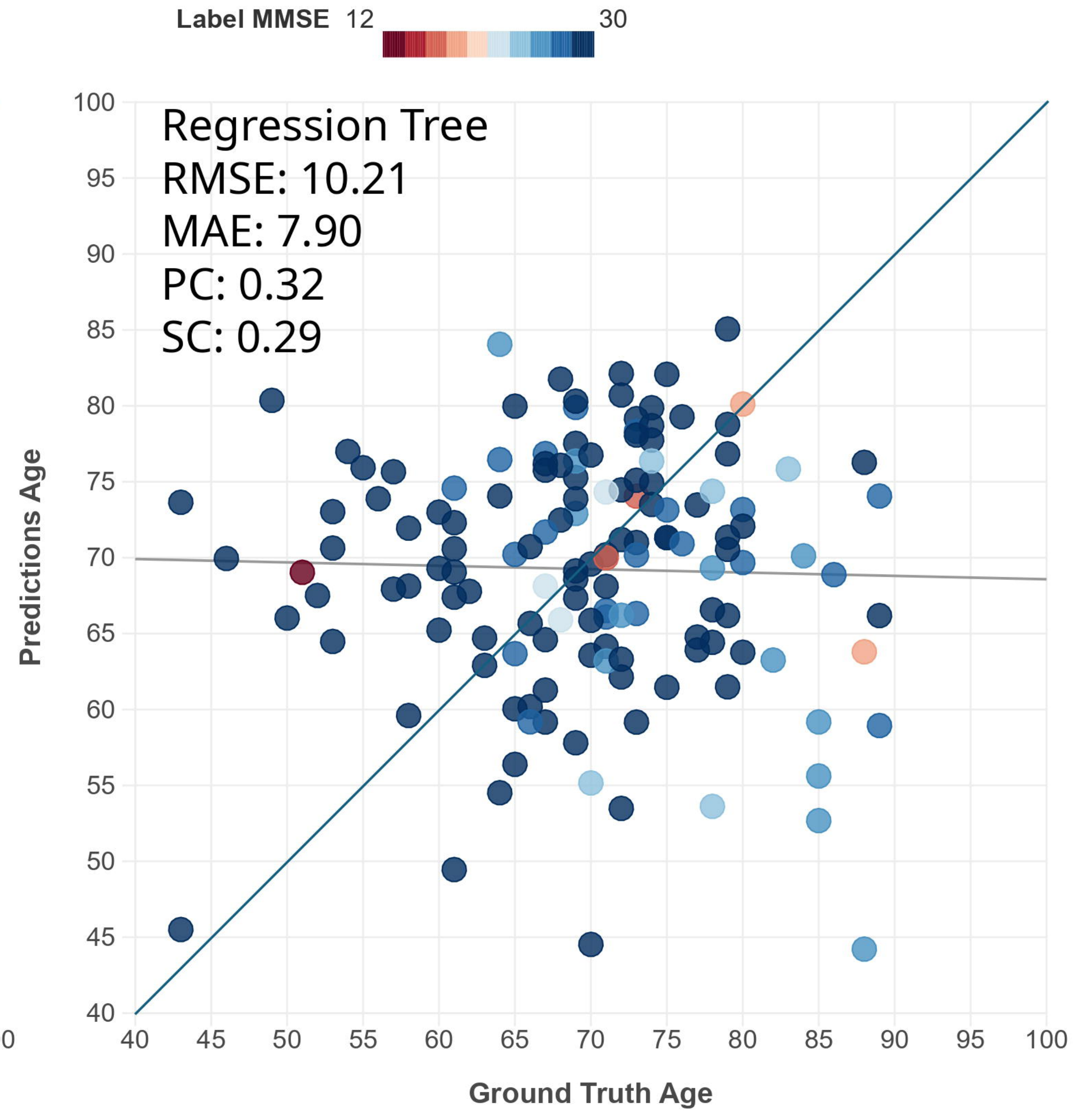

Label MMSE 12 30
SVR
RMSE: 7.56
MAE: 5.55
PC: 0.67
SC: 0.67
Predictions Age
Ground Truth Age
40 45 50 55 60 65 70 75 80 85 90 95 100

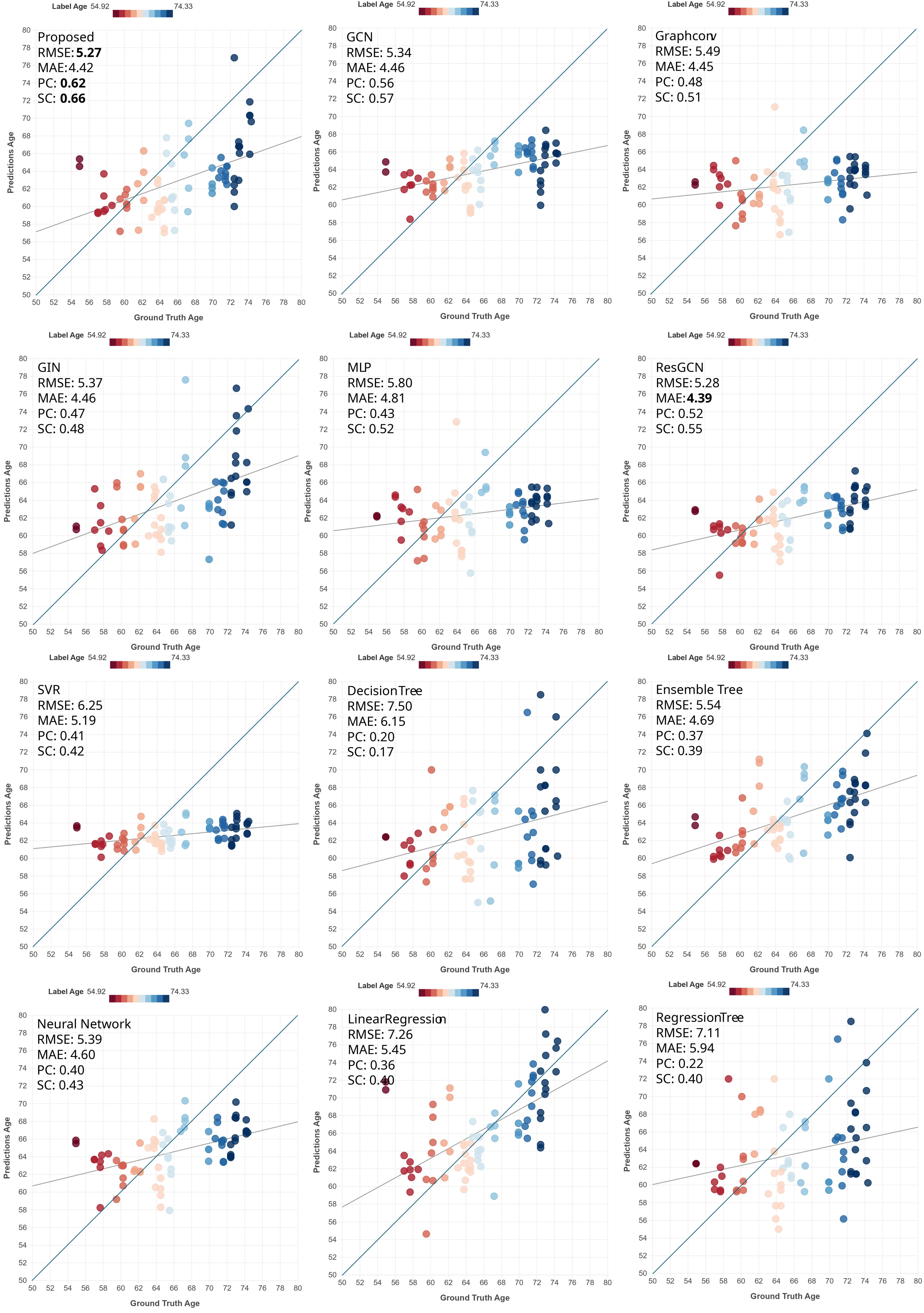

Label MMSE 12 30
Decision Tree
RMSE: 10.86
MAE: 8.37
PC: 0.29
SC: 0.28
Predictions Age
Ground Truth Age
40 45 50 55 60 65 70 75 80 85 90 95 100

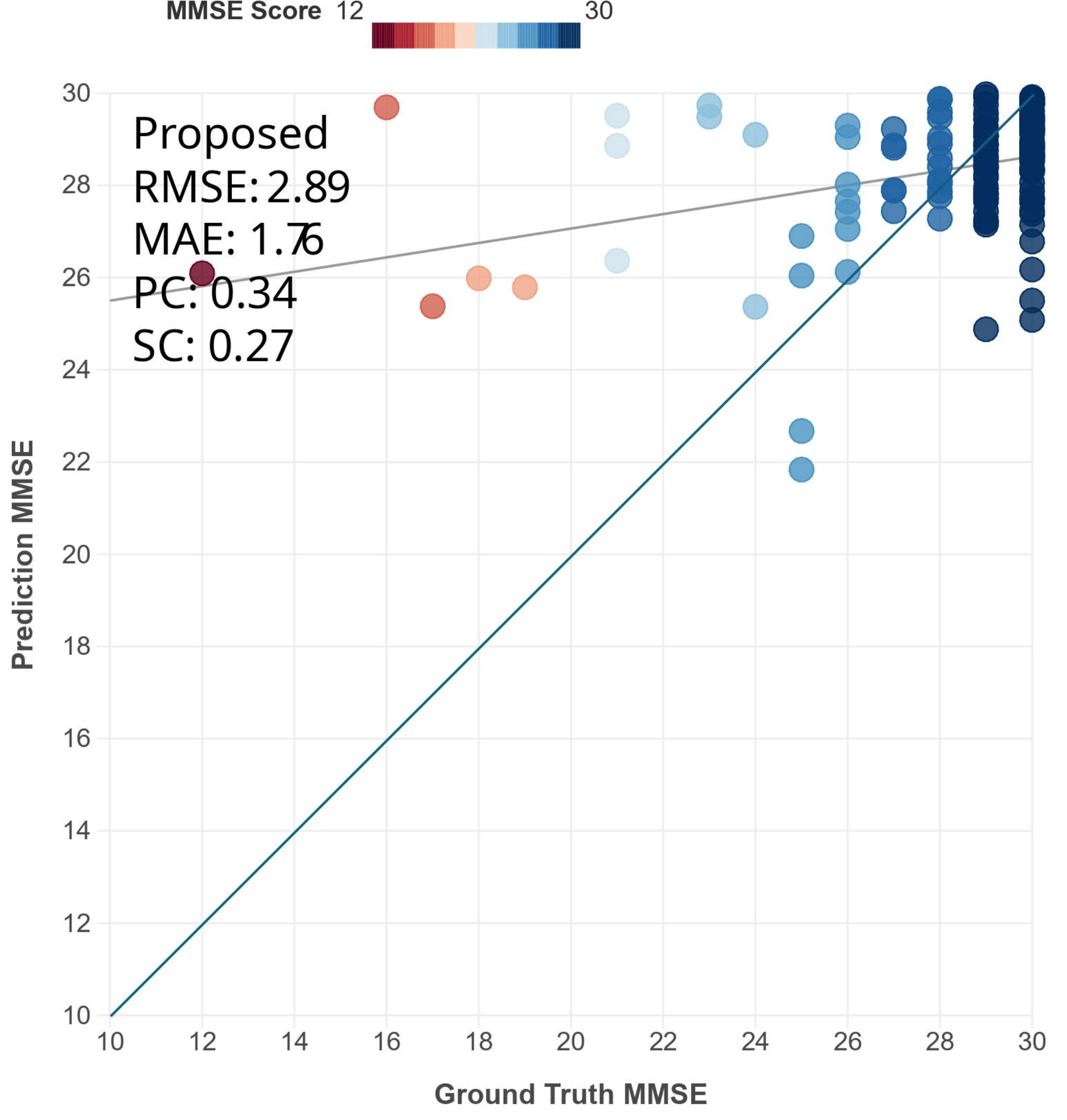

Label MMSE 12 30
Ensemble Tree
RMSE: 7.35
MAE: 5.57
PC: 0.64
SC: 0.62
Predictions Age
Ground Truth Age
40 45 50 55 60 65 70 75 80 85 90 95 100

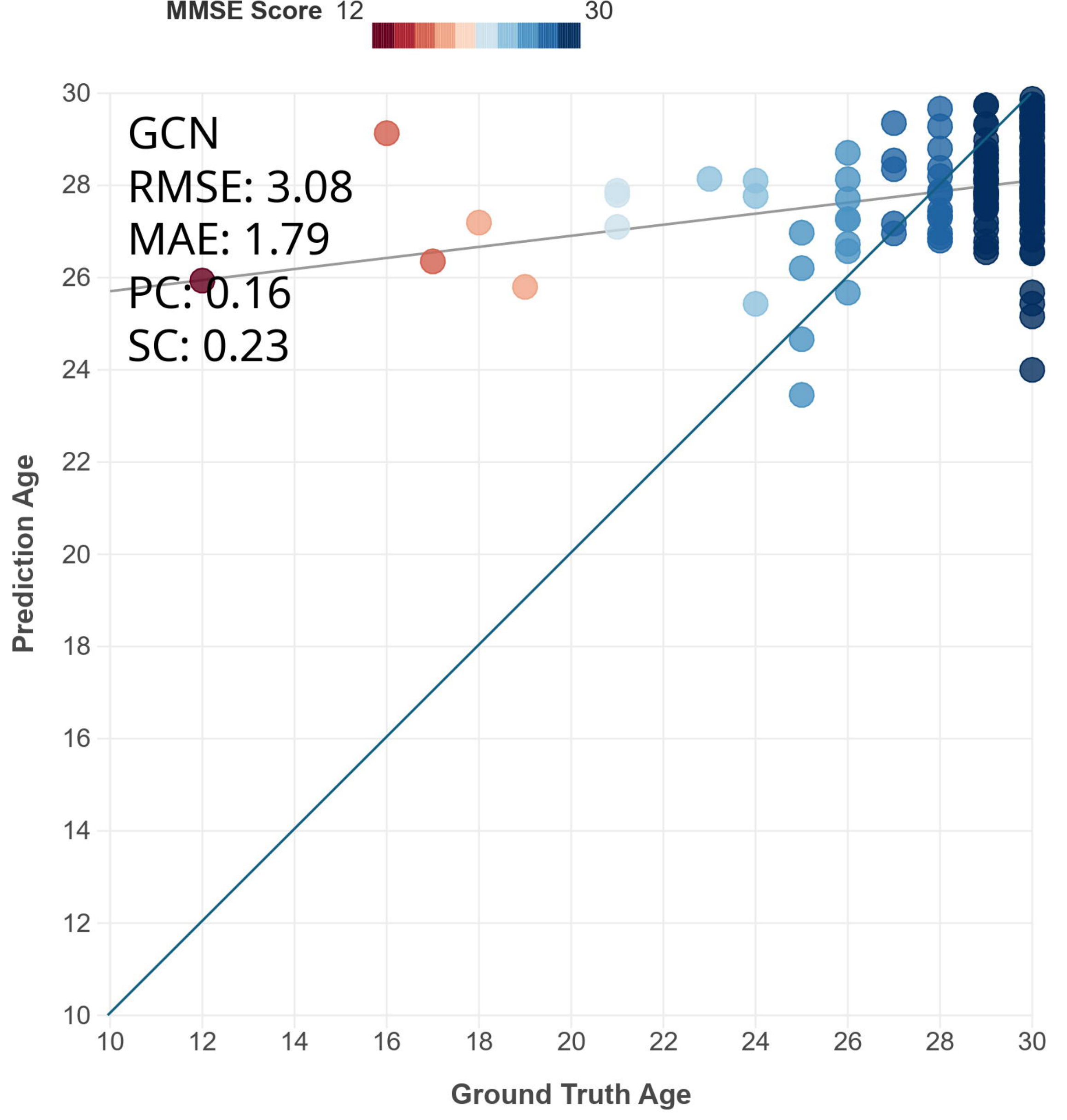

Label MMSE 12 30
Neural Network
RMSE: 7.47
MAE: 5.91
PC: 0.67
SC: 0.67
Predictions Age
Ground Truth Age
40 45 50 55 60 65 70 75 80 85 90 95 100

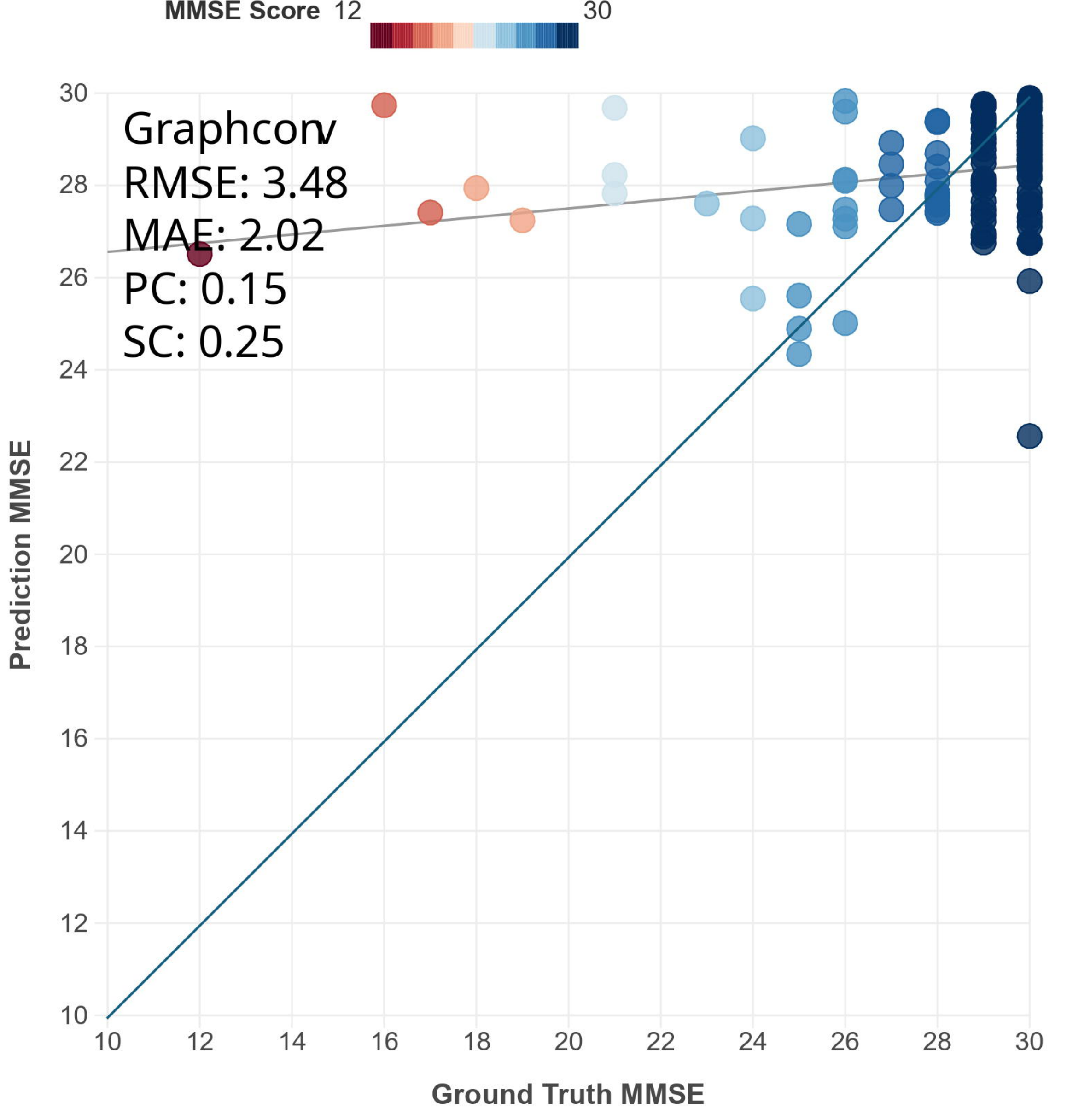

Label MMSE 12 30
Linear Regression
RMSE: 8.45
MAE: 6.57
PC: 0.65
SC: 0.66
Predictions Age
Ground Truth Age
40 45 50 55 60 65 70 75 80 85 90 95 100

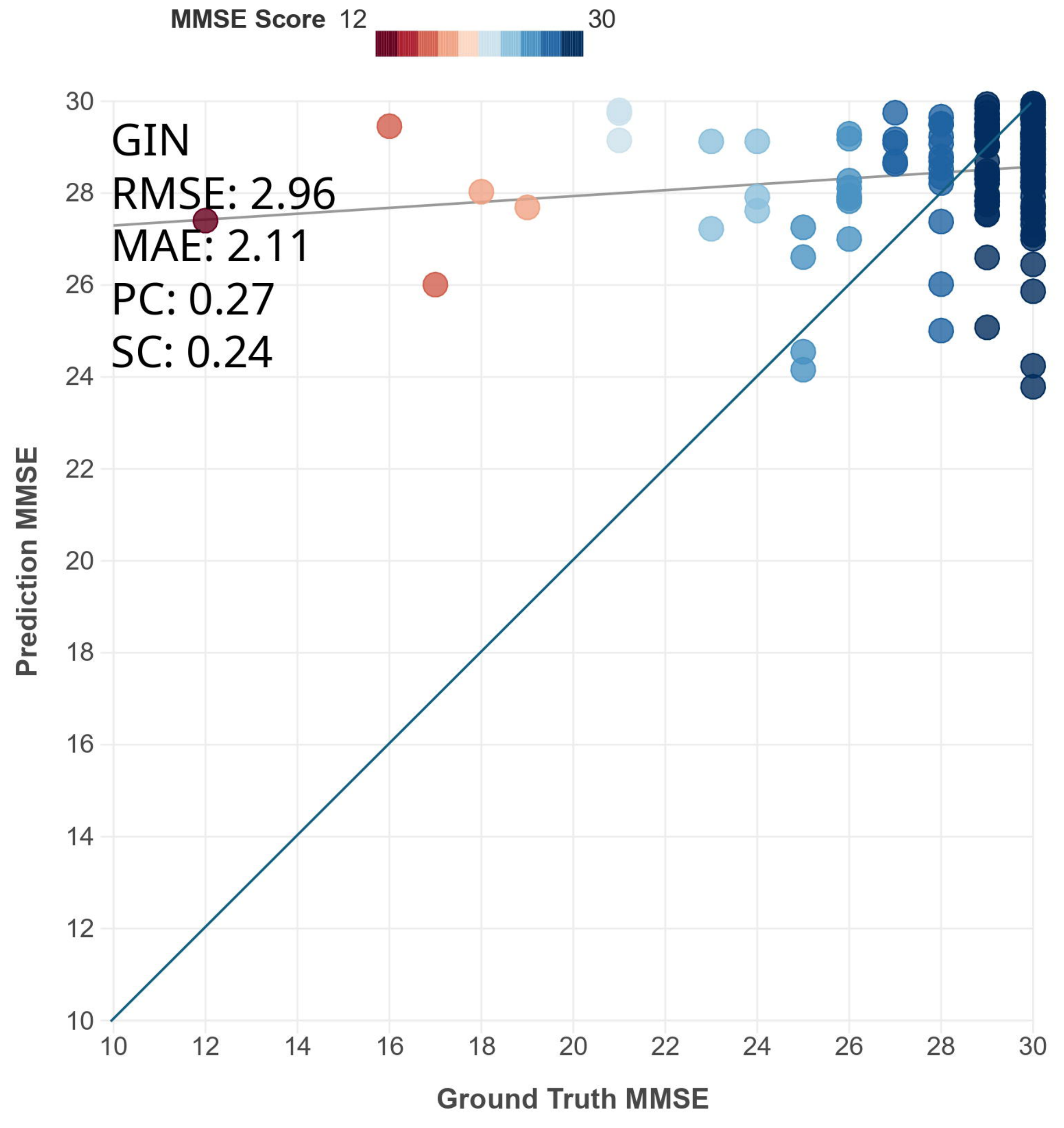

Label MMSE 12 30
Regression Tree
RMSE: 10.21
MAE: 7.90
PC: 0.32
SC: 0.29
Predictions Age
Ground Truth Age
40 45 50 55 60 65 70 75 80 85 90 95 100

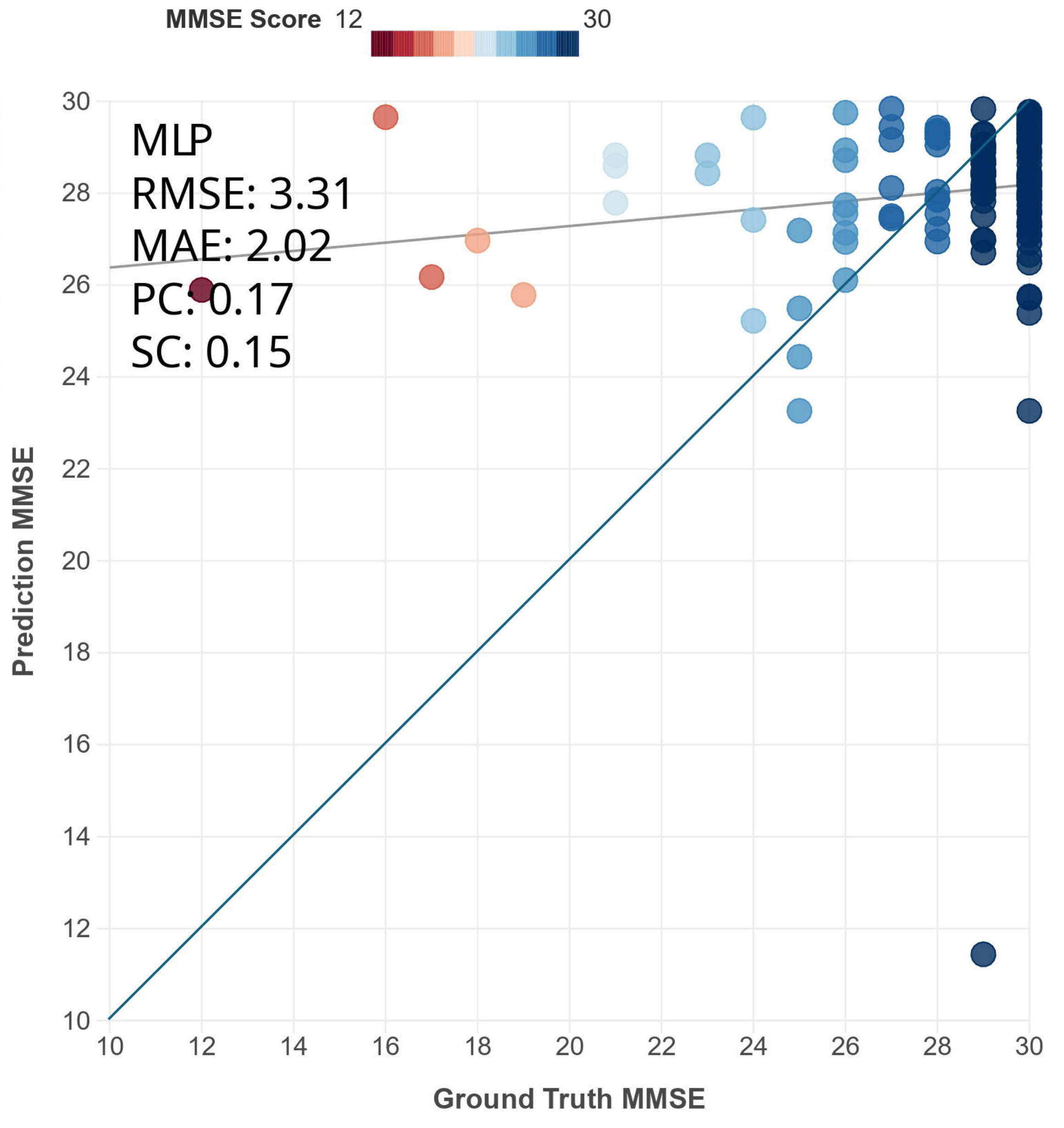

Label Age 54.92 74.33
Proposed
RMSE: **5.27**
MAE: 4.42
PC: **0.62**
SC: **0.66**
Predictions Age
Ground Truth Age
GCN
RMSE: 5.34
MAE: 4.46
PC: 0.56
SC: 0.57
Graphconv
RMSE: 5.49
MAE: 4.45
PC: 0.48
SC: 0.51
GIN
RMSE: 5.37
MAE: 4.46
PC: 0.47
SC: 0.48
MLP
RMSE: 5.80
MAE: 4.81
PC: 0.43
SC: 0.52
ResGCN
RMSE: 5.28
MAE: **4.39**
PC: 0.52
SC: 0.55
SVR
RMSE: 6.25
MAE: 5.19
PC: 0.41
SC: 0.42
DecisionTree
RMSE: 7.50
MAE: 6.15
PC: 0.20
SC: 0.17
Ensemble Tree
RMSE: 5.54
MAE: 4.69
PC: 0.37
SC: 0.39
Neural Network
RMSE: 5.39
MAE: 4.60
PC: 0.40
SC: 0.43
LinearRegression
RMSE: 7.26
MAE: 5.45
PC: 0.36
SC: 0.40
RegressionTree
RMSE: 7.11
MAE: 5.94
PC: 0.22
SC: 0.40
50 52 54 56 58 60 62 64 66 68 70 72 74 76 78 80

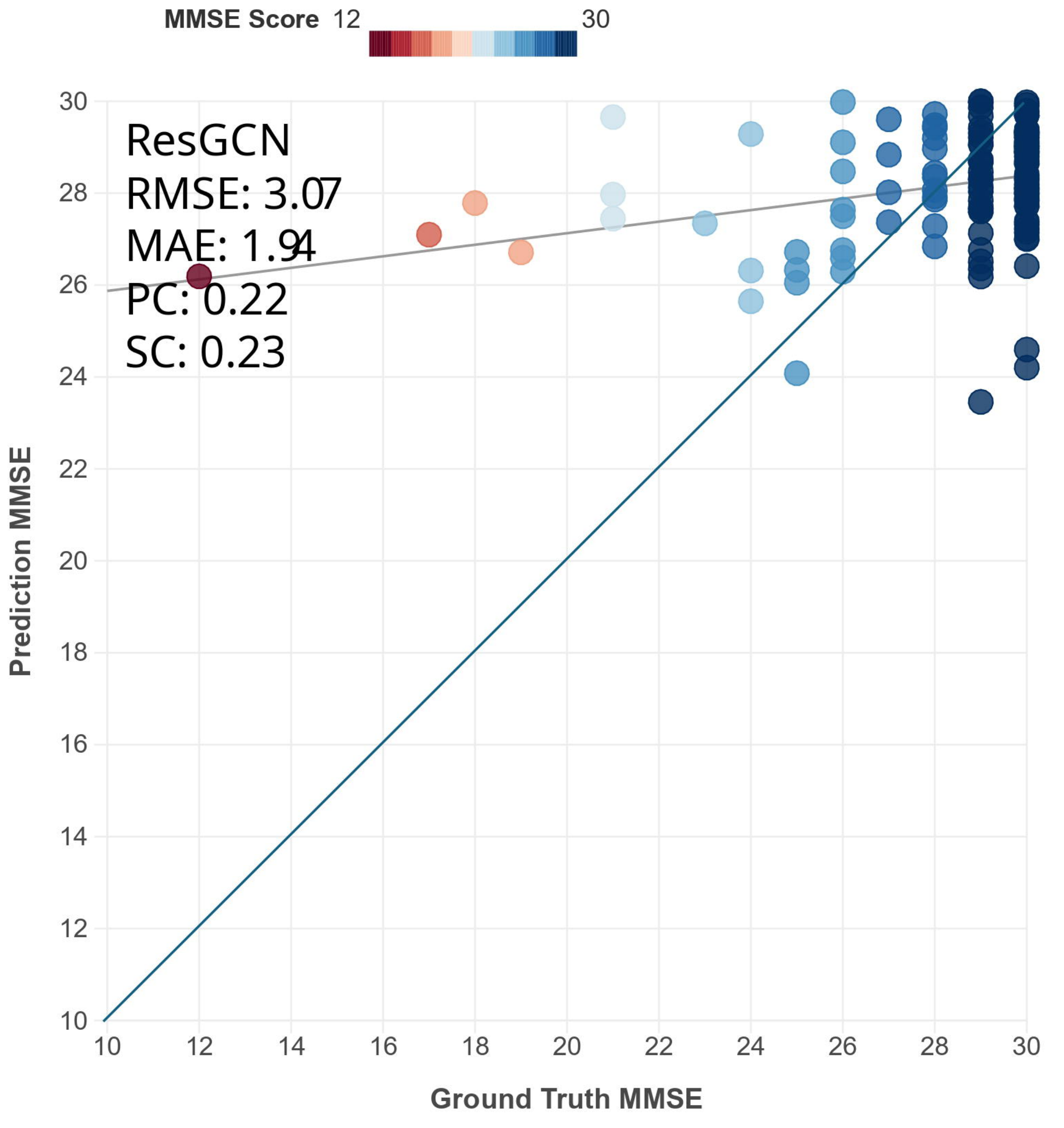

MMSE Score 12 30
Proposed
RMSE: 2.89
MAE: 1.76
PC: 0.34
SC: 0.27
Prediction MMSE
Ground Truth MMSE
10 12 14 16 18 20 22 24 26 28 30

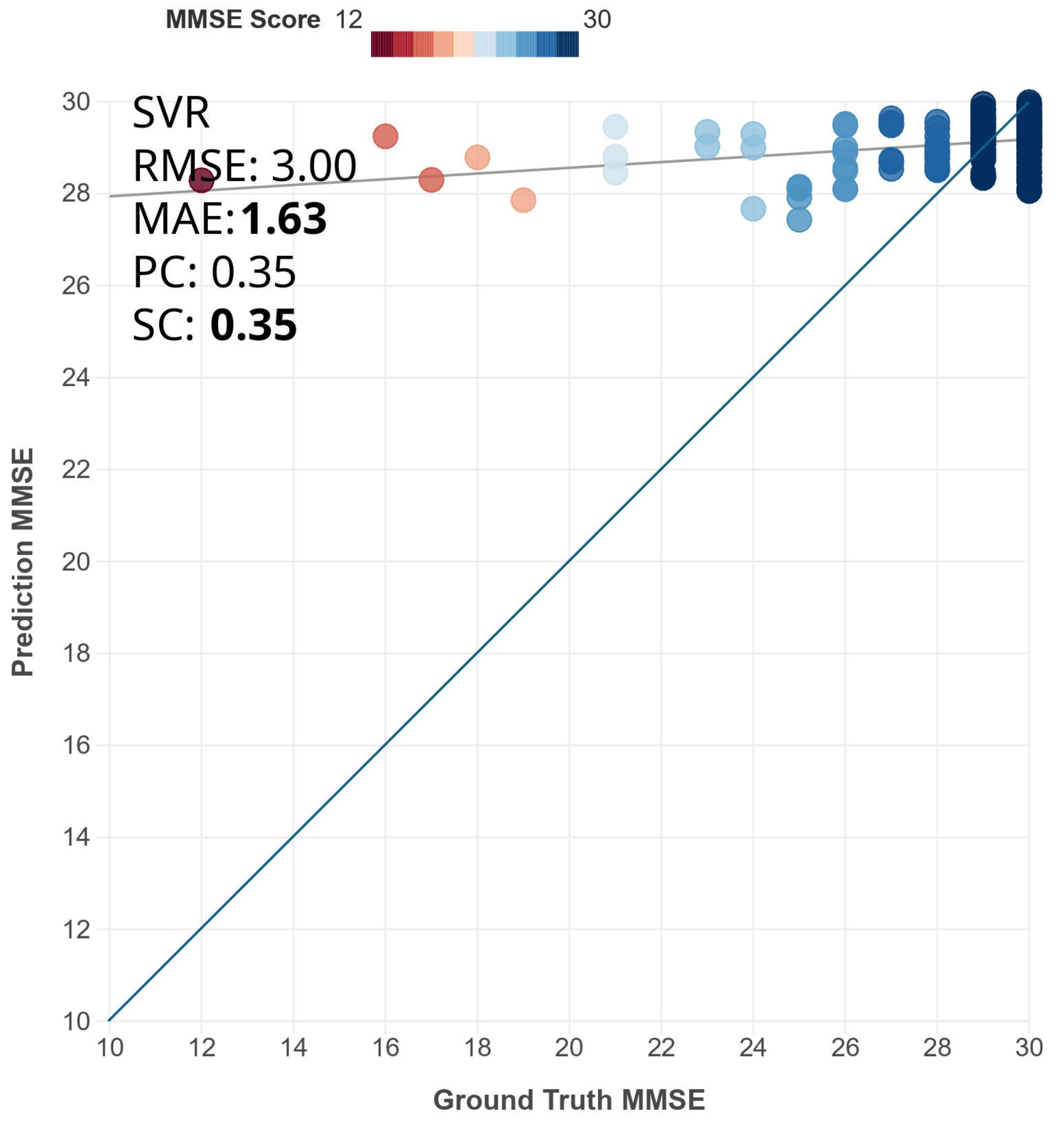

MMSE Score 12 30
GCN
RMSE: 3.08
MAE: 1.79
PC: 0.16
SC: 0.23
Prediction Age
Ground Truth Age
10 12 14 16 18 20 22 24 26 28 30

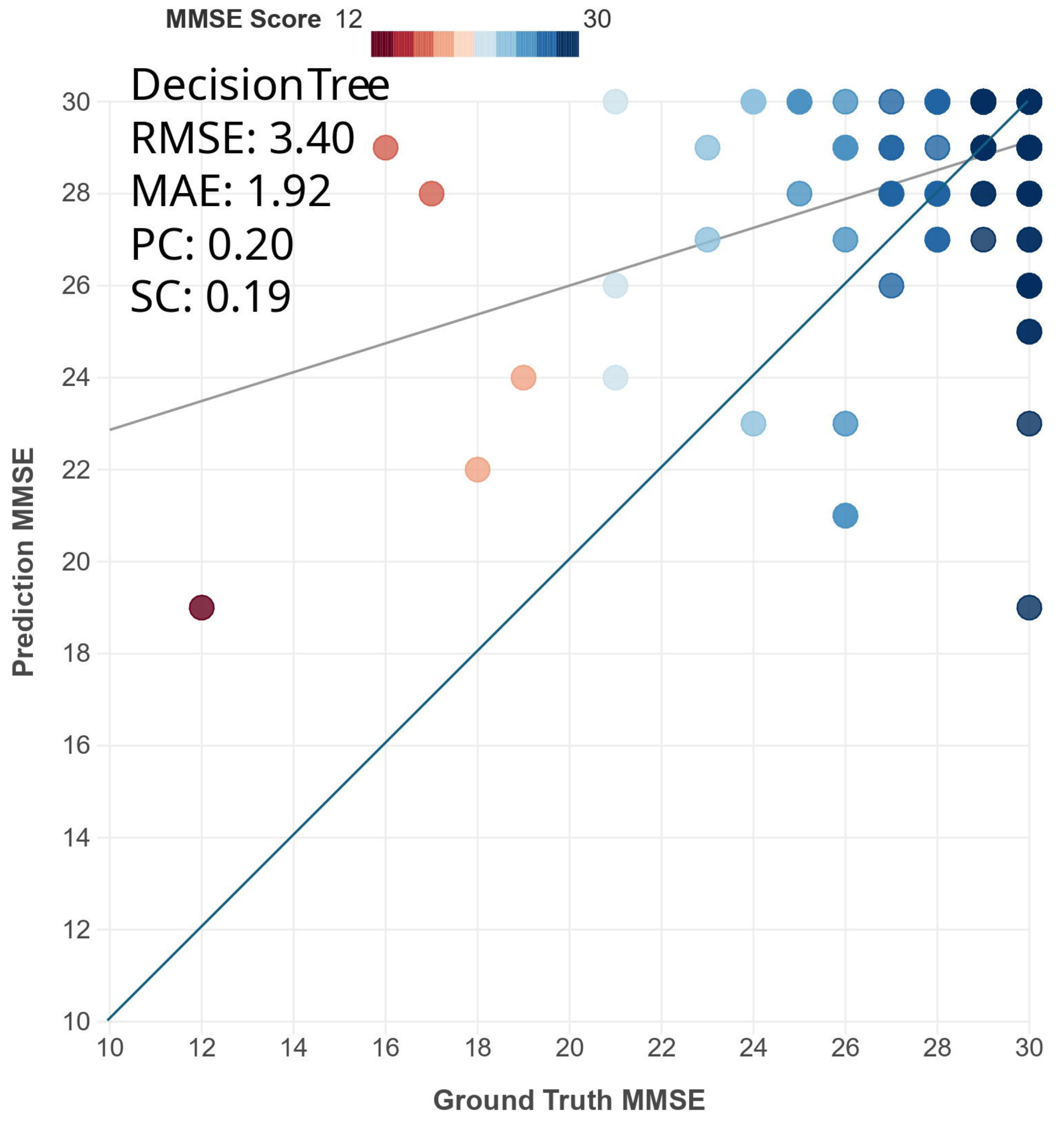

MMSE Score 12 30
Graphconv
RMSE: 3.48
MAE: 2.02
PC: 0.15
SC: 0.25
Prediction MMSE
Ground Truth MMSE
10 12 14 16 18 20 22 24 26 28 30

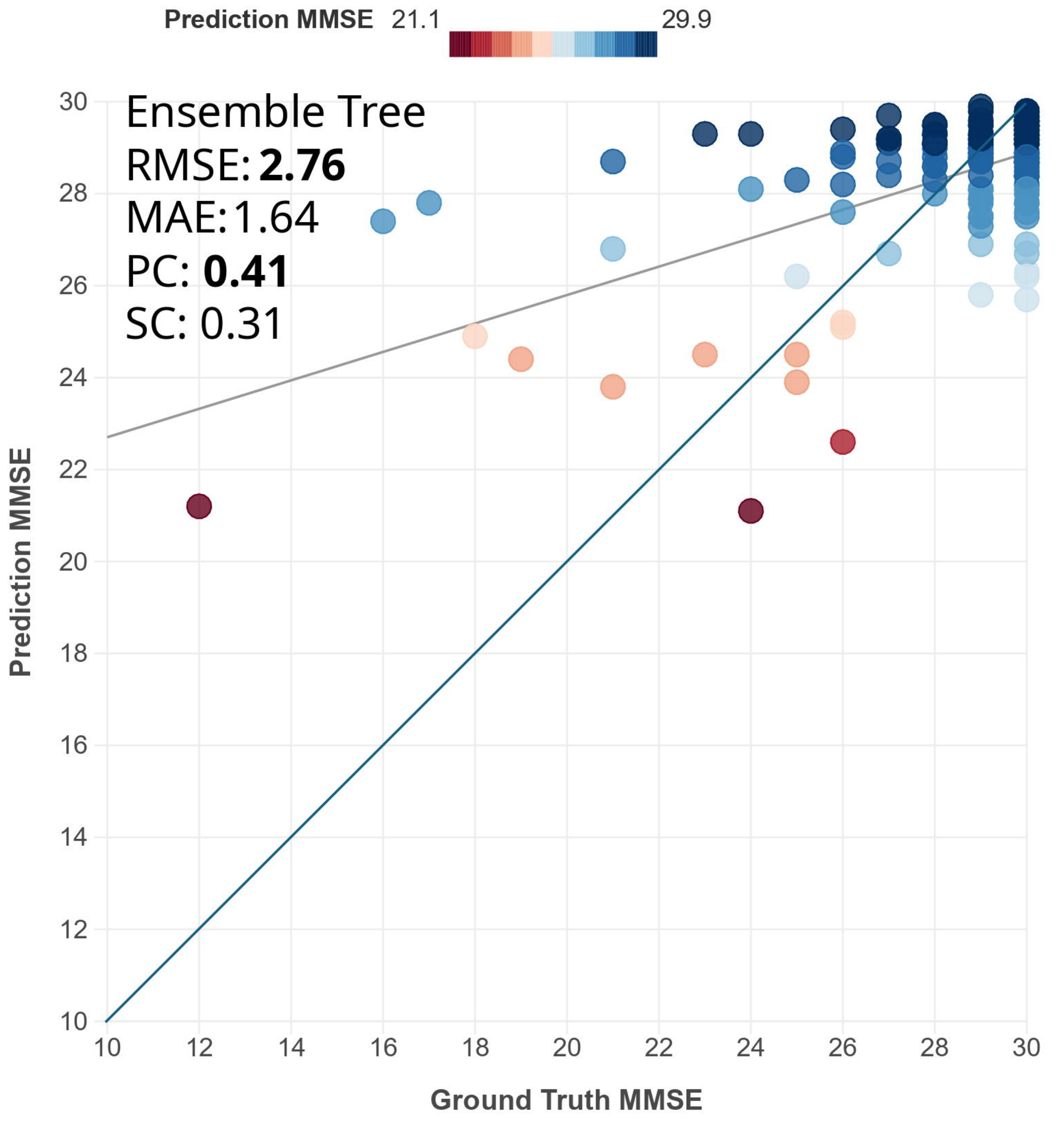

MMSE Score 12 30
GIN
RMSE: 2.96
MAE: 2.11
PC: 0.27
SC: 0.24
Prediction MMSE
Ground Truth MMSE
10 12 14 16 18 20 22 24 26 28 30

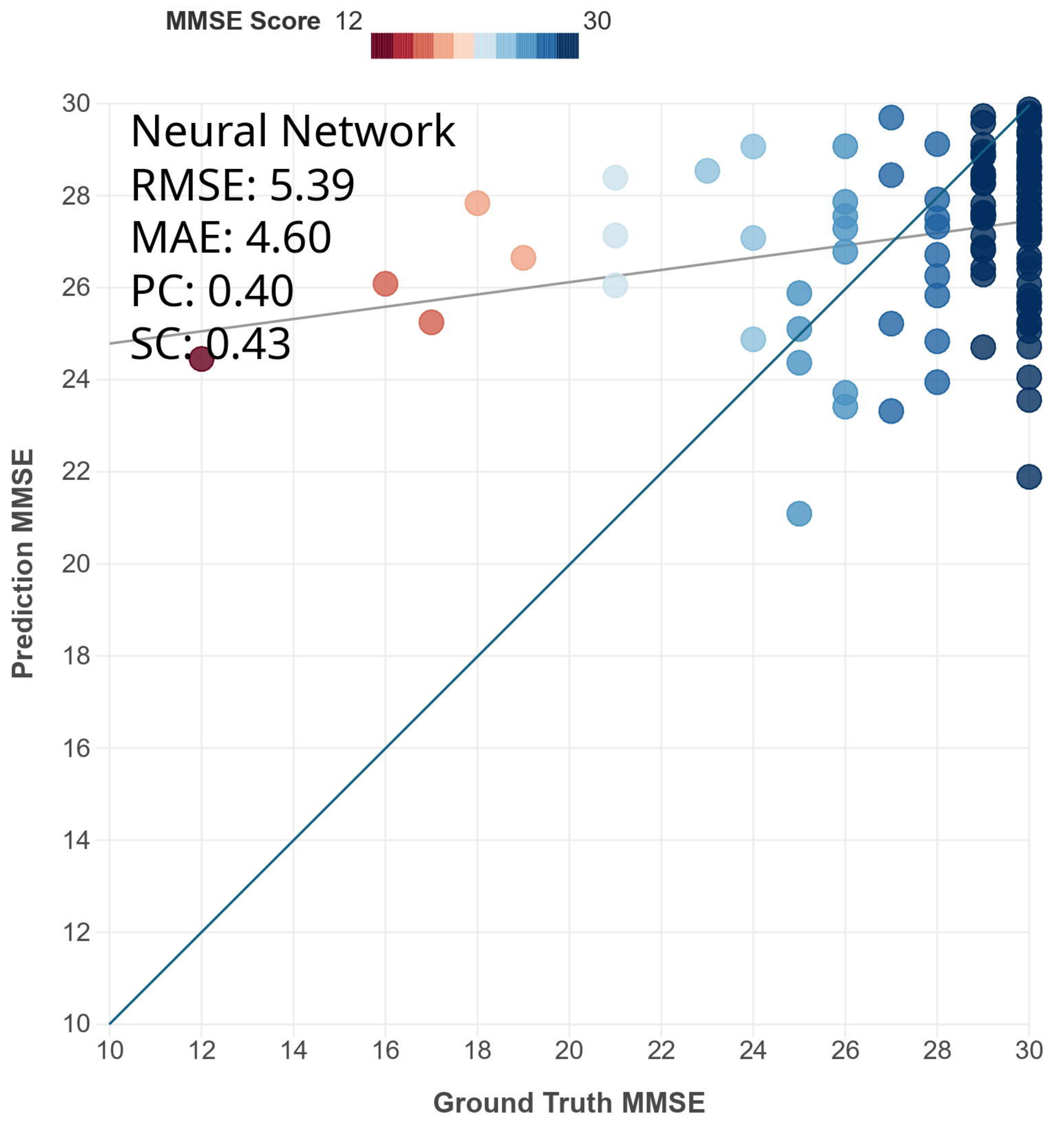

MMSE Score 12 30
MLP
RMSE: 3.31
MAE: 2.02
PC: 0.17
SC: 0.15
Prediction MMSE
Ground Truth MMSE
10 12 14 16 18 20 22 24 26 28 30

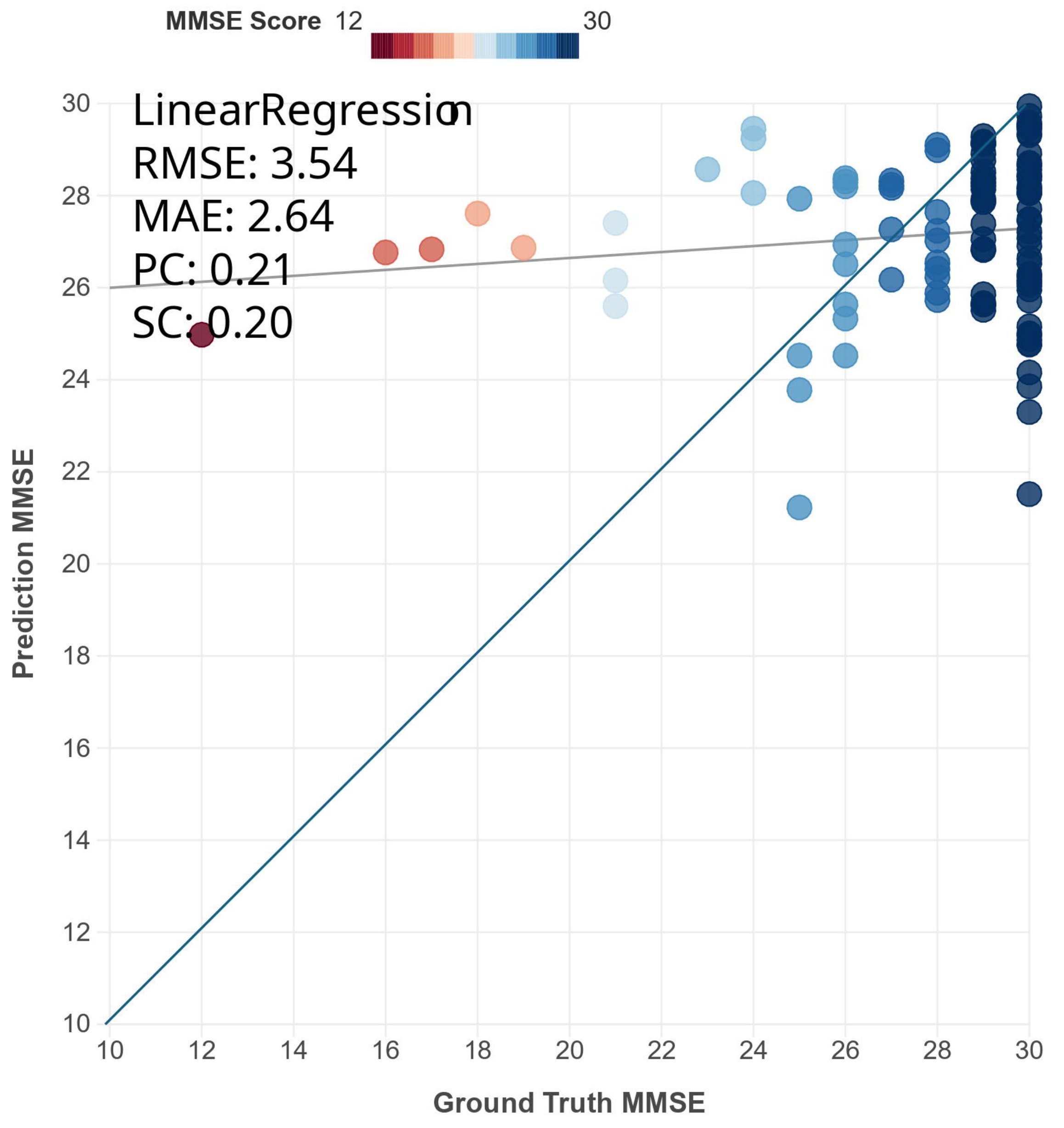

MMSE Score 12 30
ResGCN
RMSE: 3.07
MAE: 1.94
PC: 0.22
SC: 0.23
Prediction MMSE
Ground Truth MMSE
10 12 14 16 18 20 22 24 26 28 30

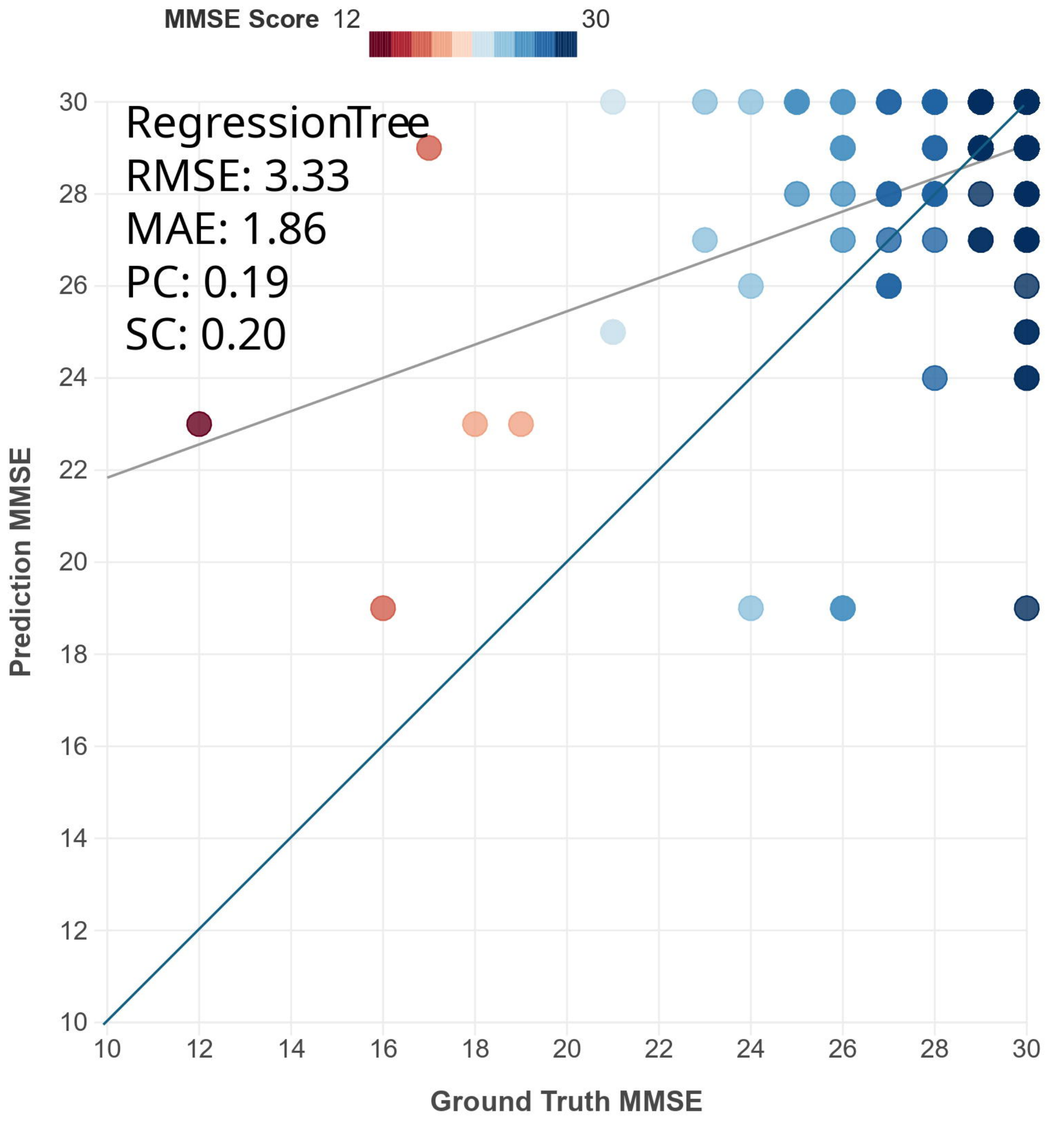

MMSE Score 12 30
SVR
RMSE: 3.00
MAE: **1.63**
PC: 0.35
SC: **0.35**
Prediction MMSE
Ground Truth MMSE
10 12 14 16 18 20 22 24 26 28 30


MMSE Score 12 30
DecisionTree
RMSE: 3.40
MAE: 1.92
PC: 0.20
SC: 0.19
Prediction MMSE
Ground Truth MMSE
10 12 14 16 18 20 22 24 26 28 30


Prediction MMSE 21.1 29.9
Ensemble Tree
RMSE: **2.76**
MAE: 1.64
PC: **0.41**
SC: 0.31
Prediction MMSE
Ground Truth MMSE
10 12 14 16 18 20 22 24 26 28 30


MMSE Score 12 30
Neural Network
RMSE: 5.39
MAE: 4.60
PC: 0.40
SC: 0.43
Prediction MMSE
Ground Truth MMSE
10 12 14 16 18 20 22 24 26 28 30


MMSE Score 12 30
LinearRegression
RMSE: 3.54
MAE: 2.64
PC: 0.21
SC: 0.20
Prediction MMSE
Ground Truth MMSE
10 12 14 16 18 20 22 24 26 28 30


MMSE Score 12 30
RegressionTree
RMSE: 3.33
MAE: 1.86
PC: 0.19
SC: 0.20
Prediction MMSE
Ground Truth MMSE
10 12 14 16 18 20 22 24 26 28 30