IBMS Wet Lab 答题总结

存D盘，记住PC ip
需要迅速下载的页面

- Week 15/ formative 1 model answer (是第二个)
- Sem2-Week5/ 20250320_IBMS3 Scientific article discussion student guide.docx (21.541 KB) 用处还行，有一个思考框架
- Sem2Week6/formative 2 model answer
- Week 8/ gene knockout mice
- Seminars

Experiment Method/ Measurement Approaches

- Only need to specific ONE method
  This step varies. Here collected the method in the previous tests:
- commercial kit for triglycerol concentration (mentioned in question)
- CRISPR-Cas9 for genetic engineering

Group Settings/ Biological Replicates

1. 检查X药物对疾病模型影响

- Experimental group: Disease model + drug
- Control Group: Disease model + placebo(安慰剂)
- Control Group baseline: Healthy model + normal process
- If possible:
  - Positive Control Group 3: Disease Model + Existing therapeutic drug

2. 确定gene X对某性状/疾病的关系

- Experimental group: inhibitor in Gene X expression (Gene knockdown: ASO in brain, RNAi, CRISPRi/ Gene Knockout: CRISPR-Cas9 System)
- Control Group: Non-targeting gRNA control - Cells receiving Cas9 plus a scrambled or non-targeting gRNA that doesn't correspond to any genomic sequence; so the cell line remains to be wild type

- (Optional) rescue group (if the edited region is in encoding region): For the rescue group, cells with Gene X knockdown/knockout will be transfected with an expression vector containing a modified Gene X sequence resistant to the targeting method (containing silent mutations), allowing restoration of functional protein expression while evading the knockdown/knockout mechanism.

3. 确定蛋白质X对于疾病的关系

For upregulated proteins in disease:

- Experimental group: Protein X knockdown/knockout (using siRNA, shRNA, CRISPR-Cas9, or specific inhibitors)
- Control group 1: Wild-type/untreated (baseline disease model)
- Control group 2: Non-targeting control (scrambled siRNA, empty vector, etc.)
- If possible: Reconstitution group (reintroducing Protein X to confirm specificity)

For downregulated proteins in disease:

- Experimental group: Protein X overexpression (using viral vectors, transgenic models, or inducible systems, like CRIPSR activator)
- Control group 1: Wild-type/untreated (baseline disease model)
- Control group 2: Empty vector control
- If possible: Dose-dependent expression group

模版补充

- 针对小鼠模型：
  - We would need a sufficiently large sample size for each group to ensure our results are statistically significant. For instance, ==10-20 mice ==per group would be reasonable, but the exact number would be determined by a power calculation. (mention sex differences if needed)
  - Ages:
    - 8-12 weeks is the most common
    - 2-6 months for mature adults
  - We shouldn't specific the time, just several weeks that follows the standard protocol
- 针对细胞模型：

    - We would need a sufficiently large sample size for each group to ensure our results are statistically significant. Typically expressed as the number of replicates (e.g., wells in a plate) per condition. At least 3-6 replicates per condition are generally recommended, but this depends on the variability of the assay. Independent experiments (biological replicates) are also crucial.
- 针对病患模型
    - We would need a sufficiently large sample size for each group to ensure our results are statistically significant. The sample size would be relative large since the greater variability in human populations. The age, health situation, ect. should be similar top reduce the influence.

Data preprocess steps for quantification

- 这一步，在答案中描述为Data transformation/ normalization step for quantification

模版

This step is necessary to ensure accurate quantification and meaningful comparison of results.

- 针对需要外部参考的：之后找到Reference standards
- 针对起始点有误差的：消除初始不同导致的影响，比如都处以每一个小组的初始值scale
- 针对测量本身有误差的：housekeep gene, internal sample(IgG)
  Each sample will be measured at least three times as technical replicates. 强调技术重复

Statistical Analysis Method

对比两组数据：

- T-test：比较两组样本均值差异，适用于正态分布数据
- Mann-Whitney U test：非参数检验，比较两组样本但不要求正态分布
- Paired-T test：比较同一样本在不同处理前后的变化
- We could use a t-test to compare the means of the two groups if the data is normally distributed and variances are equal. If not, a non-parametric test such as the Mann-Whitney U test could be used.

分类数据比较：

- Chi-square test：比较两个分类变量之间的关联性，适用于大样本量的分类数据比较
- Fisher's exact test：比较两个分类变量之间的关联性，适用于小样本量的分类数据，特别是当期望频数小于5时
- For categorical data, we would use Chi-square test for large sample sizes or Fisher's exact test when sample sizes are small or expected frequencies are low.

  多组比较：
- 单因素方差分析(ANOVA)：比较三个或更多组的均值差异
- Tukey's test：ANOVA后确定具体哪些组间存在差异
- Kruskal-Wallis test：ANOVA的非参数替代方法
- One-way ANOVA will be first applied to examine if there is any significant changes in the three/four groups; if yes, Post-Hoc test, like Tukey's test will be applied to data from the experimental group and control group

  相关性分析：
- Pearson Correlation：测量线性相关性，适用于正态分布数据
  - Pearson correlation coefficient was used to assess linear relationships between continuous variables, as data followed normal distribution
- Spearman Rank Correlation：非参数方法，测量单调关系

  - Spearman's rank correlation was employed to evaluate monotonic relationships between variables without assuming normal distribution

  生存分析
- Kaplan-Meier Curves：估计生存时间, log-rank test 显示两组之间生存的差异
  - Survival probabilities were estimated using Kaplan-Meier curves, and differences between groups were assessed with log-rank tests
- Cox Proportional Hazards Model：评估多个因素对生存的影响

  - Cox regression analysis was performed to evaluate the influence of multiple variables on survival outcomes while adjusting for confounding factors

  基因组学：
- Differential Expression Analysis：识别不同条件下差异表达的基因
  - Genes differentially expressed between conditions were identified using [method] with threshold criteria of fold change >2 and adjusted p-value <0.05
- Enrichment Analysis:：确定特定功能通路是否过度表示

    - Gene set enrichment analysis was conducted to identify biological pathways and functions overrepresented in our gene list using specific database for this task.
- Multiple Testing Correction (FDR)：控制大规模比较中的假阳性
    - To control for false discoveries in multiple comparisons, p-values were adjusted using the Benjamini-Hochberg procedure with FDR threshold of 0.05.

Ethics

1. 动物实验（Animal Experiments）
   Ethics Statement
   All animal experiments were conducted in compliance with the guidelines established by the Institutional Animal Care and Use Committee (IACUC). Procedures were designed to minimize animal suffering and distress, including the use of humane endpoints, appropriate anesthesia (e.g., isoflurane inhalation), and analgesia (e.g., buprenorphine for post-operative pain). Animal housing conditions (temperature: 22±1°C, 12h light/dark cycle) and enrichment (nesting materials, social grouping) were maintained to ensure welfare. The principles of the 3Rs (Replacement, Reduction, Refinement) were strictly followed.
2. 人类实验（Human Studies）
   Ethics Statement
   All human studies were approved by the Institutional Review Board (IRB) (Protocol No. XXX-2023) and conducted in accordance with the Declaration of Helsinki. Written informed consent was obtained from all participants prior to enrollment. For minors or vulnerable populations, consent was provided by legal guardians with additional assent from participants aged ≥12 years. Personal data were anonymized and stored securely in compliance with GDPR (General Data Protection Regulation) and local privacy laws. Risks (e.g., blood sampling) were minimized through standardized protocols, and participants were free to withdraw at any stage without penalty.
3. 细胞系实验（Cell Line Experiments）
   Ethics Statement
   All cell lines used in this study were obtained from commercial repositories (e.g., ATCC) or collaborators under material transfer agreements (MTAs). The origin and ethical sourcing of cell lines were verified:
   • Commercial cell lines: Certified by suppliers with documentation confirming

compliance with ethical standards (e.g., no use of embryonic tissues without explicit consent).
• Patient-derived cell lines: Originally collected under IRB-approved protocols (No. XXX-2010) with donor consent for research use.
Cell line identities were authenticated via STR profiling, and mycoplasma contamination was routinely tested (e.g., PCR-based assays).

4. 人原代细胞实验（Human Primary Cell Experiments）
   Ethics Statement
   Human primary cells (e.g., peripheral blood mononuclear cells, PBMCs) were isolated from donor samples collected under IRB-approved protocols. Written informed consent was obtained from all donors, specifying the scope of research use and anonymization of data. Tissue samples (e.g., surgical waste) were acquired through collaborations with certified biobanks, adhering to the Human Tissue Act (HTA) and local regulations. Experimental procedures excluded commercial exploitation of donor materials, and all data were de-identified to protect privacy
5. 人类医疗数据收集（如年龄、病史等）
   Ethics Statement
   All procedures involving the collection of human medical data (e.g., age, diagnosis, treatment history) were reviewed and approved by the Institutional Review Board (IRB) (Protocol No. XXX-2023). Written informed consent was obtained from all participants, explicitly stating the purpose of data collection, potential risks (e.g., privacy breaches), and data usage scope (e.g., academic research only). Personal identifiers (e.g., names, ID numbers) were removed, and data were pseudonymized using unique codes. Access to raw data was restricted to authorized researchers, and all records were stored in encrypted databases compliant with GDPR and HIPAA regulations.
   核心要求
   • 知情同意（明确数据用途与风险）
   • 匿名化/假名化处理
   • 数据加密与访问权限控制
6. 人类影像学数据（如MRI、CT、X光）
   Ethics Statement
   The acquisition and analysis of human imaging data (e.g., MRI, CT scans) were conducted under IRB approval (No. XXX-2024) with explicit participant consent for image storage and secondary research use. To prevent re-identification, all metadata (e.g., acquisition date, hospital codes) were removed, and facial features in head scans were anonymized using defacing

tools (e.g., Quickshear). Images were stored in a secure PACS (Picture Archiving and Communication System) with access logs audited quarterly. For public sharing, images were converted to non-DICOM formats (e.g., NIfTI) and verified against the BIDS (Brain Imaging Data Structure) standard to ensure privacy compliance.
核心要求
• 影像匿名化（去面部特征、去元数据）
• 安全存储系统（PACS/BIDS标准）
• 二次使用需重新确认伦理许可

7. 公共数据使用（如开放数据库、政府公开数据）
Ethics Statement
Public datasets (e.g., UK Biobank, NHANES) used in this study were accessed under data use agreements (DUA No. XXX-2023) and complied with the original ethical approvals granted to the data providers. All data were de-identified prior to public release, as confirmed by the source institutions. Researchers adhered to the terms of use (e.g., no attempt to re-identify individuals, non-commercial purposes) and cited data sources according to FAIR principles (Findable, Accessible, Interoperable, Reusable). Local IRB exemption (No. XXX-2023) was obtained for secondary analysis of anonymized public data.
核心要求
• 遵守原始数据提供方的伦理条款
• 禁止再识别（Re-identification）尝试
• 引用与合规声明

for monkey model

All non-human primate research will be conducted under protocols approved by the Institutional Animal Care and Use Committee (IACUC) and will comply with national and international regulations for primate research. We will implement comprehensive environmental enrichment and social housing where possible. Procedures will be designed to minimize pain and distress, with appropriate anesthesia and post-procedural analgesics. The number of animals will be limited to the minimum required for statistical validity.

Potential Nuisance Variables/ Factors or Biases

General bias

1. Clearly define the study population: Establish strict inclusion/exclusion criteria (e.g., excluding individuals with peanut allergies in a dietary study).
2. Randomization & Stratification: Use random assignment to ensure baseline characteristics (e.g., age, sex) are balanced between intervention and control groups. Apply stratification for critical variables (e.g., grouping participants by exercise experience before randomization).
3. Blinding (Masking): Single-blind: Participants are unaware of their group assignment. Double-blind: Both participants and researchers are unaware of group assignments (e.g., using coded labels instead of explicit group names).
4. Pilot testing预实验: Conduct small-scale preliminary studies to identify potential issues (e.g., participant dropout patterns or measurement inconsistencies).
5. Standardized protocols: Ensure consistency in data collection (e.g., measuring cortisol levels at fixed times of day to control for circadian rhythm effects).

For mouse model

- Potential variables could include the age and sex of the mice, their genetic background, and environmental factors such as housing conditions, diet, microbiome composition, and stress levels.
- To minimize these factors, we would use mice of the same age and sex, and from the same genetic strain.
- In addition, randomization of the mice into the control and treatment groups can also help to minimize any potential biases. Blinding of the researchers during the data collection and analysis stages can also help to reduce bias.

For cell line model

- Potential confounding variables could include differences in cell culture conditions, such as variations in media composition, temperature, or CO2 levels. To minimize these factors, we would ensure that all cell culture conditions are identical for both cell lines.
- Additional variables include passage number, cell density, mycoplasma contamination, and genetic drift over time. Regular authentication of cell lines and consistent passage protocols will be implemented.
- Batch effects in reagents and timing of experiments could introduce variability, which will be controlled by using reagents from the same lot when possible and including appropriate controls in each experimental batch.

For human patient model

- Patient demographics including age, sex, ethnicity, socioeconomic status, and comorbidities may influence outcomes and will be recorded and considered during analysis.
- Treatment history, medication use, and lifestyle factors (diet, exercise, smoking status) could confound results and will be documented.
- Selection bias may occur if recruitment strategies favor certain patient populations. We will implement broad inclusion criteria and diverse recruitment methods.
- Recall bias in patient-reported outcomes will be minimized through validated questionnaires and prospective data collection when possible.
- Loss to follow-up could skew longitudinal data and will be addressed through multiple contact methods and incentives for study completion.

Bias in methods

- CRISPR-Cas9: off-target effects could influence the results. We will use validated guide RNAs and perform off-target analysis. Control cell lines without the target modification but subjected to the CRISPR process will be included.
- Western-blotting: Antibody specificity, protein loading variations, and subjective band quantification could introduce bias. We will validate antibodies, use loading controls, and implement automated quantification software with blinded analysis.
- qPCR: Primer efficiency, reference gene stability, and threshold cycle determination can affect results. We will validate primers, use multiple reference genes, and apply consistent threshold settings across all samples.
- RNA-seq: Library preparation biases, batch effects, and computational analysis choices can influence differential expression results. We will prepare libraries simultaneously, include spike-in controls, and apply multiple normalization methods to ensure robustness.
- Flow cytometry: Gating strategies, fluorophore compensation, and instrument calibration can introduce variability. We will use fluorescence-minus-one controls, consistent gating protocols, and regular instrument quality control.
- Microscopy: Selection bias in field of view, inconsistent exposure settings, and subjective image analysis can skew results. We will implement systematic

random sampling, standardized acquisition parameters, and automated quantification algorithms.

- Animal behavior tests: Observer presence, time of day, and environmental conditions can affect behavioral outcomes. We will conduct tests at consistent times, use automated tracking systems when possible, and maintain consistent environmental conditions.
- Mass spectrometry: Sample preparation variability, ion suppression, and peak identification can bias metabolomic or proteomic analyses. We will include internal standards, run quality control samples, and use multiple technical replicates.
- ELISA: Plate position effects, edge effects, and inconsistent washing can lead to systematic errors. We will randomize sample positions, include standard curves on each plate, and use automated washing systems when available.
- ChIP-seq: Antibody specificity, chromatin fragmentation variability, and PCR amplification biases can affect results. We will validate antibodies using knockout controls, optimize sonication protocols, and minimize PCR cycles.
- Single-cell analyses: Cell isolation methods, doublet contamination, and computational clustering algorithms can bias population identification. We will implement doublet removal strategies, include control populations, and apply multiple clustering approaches.
- Histology: Tissue processing artifacts, staining variability, and region selection can skew quantification. We will process all samples simultaneously, include positive and negative controls, and implement systematic sampling of tissue sections.
- Bioinformatic analyses: Algorithm selection, parameter settings, and reference database choices can significantly impact results. We will apply multiple analytical approaches, perform sensitivity analyses with different parameters, and validate key findings using alternative methods.
- Metabolic assays: Circadian rhythm effects, feeding status, and stress responses can influence metabolic measurements. We will standardize testing times, fasting conditions, and acclimation periods prior to measurements.
- Transient transfection: Variable transfection efficiency between samples can affect protein expression levels. We will include transfection controls, optimize conditions, and normalize results to transfection efficiency.
- Luciferase assays: Substrate degradation, cell number variations, and instrument sensitivity can bias reporter activity measurements. We will use

fresh substrate, normalize to cell number or protein content, and include internal control reporters.

- Co-immunoprecipitation: Non-specific binding, antibody cross-reactivity, and wash stringency can affect protein interaction results. We will include IgG controls, validate antibodies, and optimize wash conditions for each interaction.
- Organoid cultures: Batch-to-batch matrix variability, size heterogeneity, and differentiation state can influence outcomes. We will use matrix from the same lot, standardize seeding density, and characterize differentiation markers.
- Calcium imaging: Dye loading variability, photobleaching, and motion artifacts can bias signal quantification. We will normalize to baseline fluorescence, correct for bleaching, and implement motion correction algorithms.
- Electrophysiology: Electrode placement, seal quality, and cell health can affect recording quality. We will establish minimum seal resistance criteria, monitor access resistance, and exclude recordings with unstable baselines.
- In vivo imaging: Anesthesia effects, tissue depth, and motion artifacts can influence signal detection. We will standardize anesthesia protocols, implement depth correction, and use respiratory gating when applicable.
- Pharmacological studies: Drug solubility, stability, and off-target effects can confound interpretations. We will prepare fresh solutions, verify activity, and use multiple structurally distinct compounds targeting the same pathway.

How the tutorial ask us to think...

- 首先，Hypothesis/ Experimental purpose and question?
  - Model设计相关--> Model system是什么？Cells/ Animals/ human population? Why Do they chose the model? Are there ethic to be considered?
    - --> Group setting:
      - Negative control?
      - Positive control?
      - Internal Control for normalisation?
      - biological replicates?
      - How to control the variables/ factors?

    - --> Measurement approaches是什么？PCR, WB, IF, IHC, Sequencing....?
        - Are there technical replicates?
        - Why to choose this approach?
    - 结果相关--> Experimental output: image/ counts/ numbers
        - Data process?
        - Quantitative or qualitative?
        - Statistical approaches?
- Pilot experiment: (Critical thinking part)
    - 是指在进行大规模试验前进行的先导的小规模的测试试验，验证可行性
    - Can the data be reproduced independently?
    - Can they reproduce previous findings?
    - Awareness of potential bias?
    - Can the results be validated by an alternative approach?
    - How would you design it?