**康复项目数据与描述**

**项目名称**：机构护理和强化训练康复与标准监督建议康复的比较。

**目标人群**：经历过心肌梗塞的男性患者。这些患者主要是因为状况不佳和活动迟缓而被推荐进行康复。

这类患者倾向于保持被动，这会恶化本已不佳的预后。因此，开发了一种新的、更积极的方法。新方法的主要目标是鼓励患者增加活动量。患者将接受一个包含强化训练和心理治疗的方案。

为了评估这种新的治疗方法并与标准治疗进行比较，符合纳排标准的患者将被纳入研究，患者被随机分配到**标准治疗**或**强化治疗组**。两组患者的治疗时间均为3个月，患者被要求在6个月后进行随访，以评估治疗的长期效果。

该研究的**主要结果指标**是工作能力（单位：瓦特）。指标在基线、3个月和6个月后进行测量。此外，如果治疗3个月后工作能力至少增加了30瓦，治疗被视为成功（分类变量SUCC3，并用作**次要结果变量**）。其他次要的结果指标包括胆固醇、甘油三酯和心理状态的变化。

所有的变量列表如下：

**变量列表**

|  |  |  |  |
| --- | --- | --- | --- |
| **NR** | **Name** | **Description** | **Code/evels** |
| V1 | PAT | Patient identification |  |
| V2 | DATE | Date of entry (year,month) |  |
| V3 | TREAT | Type of treatment/treatment group | 0=Standard treatment 1=Intensive treatment |
| V4 | AGE | Patient age in years |  |
| V5 | SMOKE | Smoker | 0=No,1=Yes |
| V6 | HYPERTEN | Hypertension | 0=No,1=Yes |
| V7 | DIAB | Diabetes | 0=No,1=Yes |
| V8 | CLAUD | Claudication | 0=No,1=Yes |
| V9 | B\_BLOCK | Beta-blocker | 0=No,1=Yes |
| V10 | CAUSE | The cause of admission | 1=angina 2=dyspnea  3=anxiety, depression 4=chest pain |
| V11 | WE | Maximum work capacity at entry/baseline |  |
| V12 | WD | Maximum work capacity at discharge (3 months of treatment) |  |
| V13 | WEFF | Change in work capacity between entry and discharge, calculated as WD-WE |  |
| V14 | W6 | Maximum work capacity at 6 months follow-up |  |
| V15 | WEIGHTE | Weight at entry |  |
| V16 | WEIGHTD | Weight at discharge |  |
| V17 | CHOLE | Cholesterol at entry |  |
| V18 | CHOLD | Cholesterol at discharge |  |
| V19 | TRIGE | Triglycerides at entry |  |
| V20 | TRIGD | Triglycerides at discharge |  |
| V21 | PSYCHE | Psychological status at entry | 1=bad  .  .  7=good |
| V22 | PSYCHD | Psychological status at discharge | 1=bad  .  .  7=good |
| V23 | SUCC3 | Success, an effect≥30 Watt after 3 month of treatment. | 0=failure 1=success |
| V24 | SUCC6 | Success, an effect≥30 Watt after 6 month of treatment. | 0=failure 1=success |
| V25 | SALARY | Patients monthly salary in SEK |  |
| V26 | TRAINING TIME | Average training time in hours/day |  |
| V27 | COST | Treatment cost in SEK |  |

**一、应用问题：描述性统计（32分）**

1. **为了比较两组的工作能力，请制作在入组时（WE）、出院时（WD）以及6个月时（W6）的箱线图。（2分）**
2. **下列哪项是错误的？（2分）**

A. 与强化组相比，标准组在变量WE的入组时有更多的异常值。

B. 箱线图显示，强化组在变量WD的非异常值范围似乎比标准组更宽。

C. 在治疗期间达到的工作能力，在出院后对两组都能维持。

D. 以上都不是。

1. **填写两个表格，用描述性统计数据，并回答下面的问题。（2分）**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **AGE** | N | MEAN | SD | MIN | P25 | MEDIAN | P75 | MAX | IQR |
| All patient's |  |  |  |  |  |  |  |  |  |
| Standard group |  |  |  |  |  |  |  |  |  |
| Intensive group |  |  |  |  |  |  |  |  |  |

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **WE** | N | MEAN | SD | MIN | P25 | MEDIAN | P75 | MAX | IQR |
| All patient's |  |  |  |  |  |  |  |  |  |
| Standard group |  |  |  |  |  |  |  |  |  |
| Intensive group |  |  |  |  |  |  |  |  |  |

1. **圈出正确答案 （2分）**

A. 标准组的患者比强化组的患者年龄稍大。

B. 入组时的工作能力，强化组比标准组高。

C. 入组时的最小工作能力，强化组比标准组高。

D. 全部。

1. **查看两组受试者吸烟、患糖尿病和使用β受体阻滞剂的比例。（2分）**

|  |  |  |
| --- | --- | --- |
|  | Proportions in Standard group | Proportions in Intensive  group |
| Smokers (smoke=1) |  |  |
| Diabetes (diab=1) |  |  |
| Beta blockers (b\_block=1) |  |  |

1. **圈出正确答案（可以使用四舍五入的值） （2分）**

A. 标准组中糖尿病患者的比例是9%。

B. 标准组中吸烟者的比例是89%。

C. 标准组吸烟者更多，β受体阻滞剂使用者比强化组少。

D. B和C。

1. **对于40岁以下的人来说，得心肌梗塞是不寻常的。计算样本中40岁以下患者的数量和相应比例。圈出正确答案 （2分）**

A. 40岁以下患者数量为8，40岁以下患者比例为4%。

B. 40岁以下患者数量为11，40岁以下患者比例为5.5%。

C. 40岁以下患者数量为192，40岁以下患者比例为96%。

D. 以上都不是。

1. **创建一个名为CWC（工作能力变化）的新变量，工作能力变化是WD和WE之间的差值。并制作按TREAT分类的干预成本（COST）和CWC的散点图。（2分）**
2. **下列哪项不是真的（2分）**

A. 对于强化组，随着CWC的增加，COST也增加。

B. 散点图显示了干预成本（COST）和工作能力变化（CWC）之间的关系，针对两种治疗组。

C. 对于两组来说，COST和CWC之间存在正相关。

D. 使用标准治疗，实现最高和最低CWC的成本相同。

**主要结果变量WEFF与年龄的关系**

1. **制作按两治疗组分类的WEFF与年龄的散点图。（2分）**
2. **圈出错误陈述 （2分）**

A. 散点图显示了年龄和WEFF之间的关系。

B. 对于两组来说，似乎年龄和WEFF之间存在负相关。

C. 随着年龄的增长，WEFF减少。

D. 以上都不是。

**为两组治疗的患者制作TRIGE和WEFF变量的分布图**

1. **给不同治疗组制作TRIGE变量、WEFF变量的直方图。（2分）**
2. **关于变量的分布，下列哪项是正确的 （2分）**

A. 标准组的TRIGE变量的直方图似乎比强化组更向右偏。

B. 标准组的WEFF变量的直方图是偏态的。

C. 强化组的TRIGE变量的直方图似乎比标准组更不对称。

D. 两组的WEFF变量的直方图都是偏态的。

**对于次要结局指标心理状态：**PSYCHE**（入组时）、**PSYCHD**（出院时）：**

1. **比较两组治疗组的心理状态，检查入组时和出院时的心理状态的中位数和百分比。下面哪个说法是不对的：（2分）**

A. 与强化组相比，标准组在入组时的中位数值较高。

B. 如果我们将变量PSYCHE变成二分类变量，规定1-3为“差”以及5-7为“好”，46%的标准组和61%的强化组的病人在入组时将被分类为 “差”。

C. 如果我们将变量PSYCHD变成二分类变量，规定1-3为“差”以及5-7为“好”，11.5%的标准组和2.9%的强化组的病人在入组时将被分类为“差”。

**D. 与标准组相比，强化组在出院时的中位数值较高。**

1. **分组为入组时和出院时的心理状态制作条形图。（2分）**
2. **下列哪项是不正确的： （2分）**

A. 条形图显示，与强化组相比，标准组在入组时心理状态为5的受试者百分比更高。

B. 与标准组相比，强化组在出院时心理状态最高的患者比例（状态=7）更高。

C. 条形图表明，强化组在出院时没有人的状态为1和2。

D. 以上都是。

**二、应用问题：差异比较（60分）**

**拟比较两组的工作能力变化情况，请回答17~21的问题：**

1. **强化组的工作能力，比较在治疗前（WE）和治疗后（WD）是否存在差异，请用规范的表格及统计学语言给出您的结论。（5分）**
2. **对照组的工作能力，比较在治疗前（WE）和治疗后（WD）是否存在差异，请用规范的表格及统计学语言给出您的结论。（5分）**
3. **治疗后，拟比较两组的工作能力的变化（WEFF）是否有差异，请为该研究问题建立假设检验步骤。并用规范的表格及统计学语言给出您的结论。（5分）**

**拟比较两组治疗的成功率（SUCC3），请回答22~26的问题：**

1. **绘制治疗组别、成功率的四格表。（5分）**
2. **对照组以及强化组的成功率分别是多少？请绘制规范的统计学表格。（5分）**
3. **对照组以及强化组的成功率是否有差异？请为该研究问题建立假设检验步骤。并请用规范的表格及统计学语言给出您的结论。（5分）**

**三、应用问题：回归分析（38分）**

**模型1：康复研究的目的是评估对经历过心肌梗塞的患者的治疗效应。回想一下t检验的结果。t检验是一种分析，只能考虑一个二元组变量。假设有其他重要的解释变量，你希望在分析中进行调整。回想一下你在描述性统计应用练习中制作的分类散点图。年龄和WEFF之间是否存在关联，对两组治疗组来说？（问题1-4）**

1. **如果你的目标是想知道WEFF随年龄增长的的系数，以下哪种分析适合评估年龄和WEFF之间的关系？（2分）**

A. 相关性

B. T检验

C. 简单线性回归

D. 卡方检验

1. **用散点图检查年龄和WEFF之间的关系。将结果变量放在y轴上，解释变量放在X轴上。你能得出什么结论？ （2分）**

A. 患者年龄和WEFF之间存在负线性关系。

B. 患者年龄和WEFF之间存在正线性关系。

C. 由于变异性太大，年龄和WEFF之间没有明显的关系。

D. 患者年龄和WEFF之间的关系不是线性的。

1. **执行一个简单线性回归分析，以WEFF作为结果变量，年龄作为预测变量，下列哪项陈述是真实的？ （2分）**

A. 斜率估计为-2.00

B. 斜率估计为-11.76

C. 斜率估计为-0.64

D. 斜率是3.56

1. **模型解释的结果变量方差的分数是多少？ （2分）**

A. 100% B. 64.1% C. 40.9% D. 41.1%

**模型2：从模型1的分析中，我们可以确定患者年龄对工作能力有一些影响。研究人员还希望进一步研究β受体阻滞剂、吸烟习惯和糖尿病诊断对结果的影响。建立一个回归模型，评估矫正患者年龄、β受体阻滞剂、吸烟习惯和糖尿病诊断后的治疗效应差异。考虑一个可能“过拟合”的初步模型，包括患者的治疗变量和其它需要矫正的变量。**

1. **调整其他解释变量后，两组治疗之间的工作能力变化的估计差异是多少？ （2分）**

A. 0.138 B. 3.56 C. 9.00 D. 无法得知

1. **通过调整年龄、β受体阻滞剂、吸烟和糖尿病，治疗效应差异的估计如何受到影响？与使用两个样本t检验时的结果进行比较。 （2分）**

A. 调整模型中治疗效应的差异增加了

B. 调整和未调整模型之间的治疗效应差异相似

C. 调整模型中治疗效应的差异减少了50%以上

D. 调整后的估计治疗效应差异不可解释，因为它在统计上不显著。

1. **对未调整和调整后的治疗效应对结果（工作能力变化）的影响之间的差异，可能的解释是什么？ （2分）**

A. 未调整模型中有影响的观察现在影响较小

B. 模型II可能被过度指定，因此估计可能不正确

C. 添加的解释变量之一可能是治疗和结果之间关系的可能混杂因素

D. 以上都对

1. **模型2可以解释的工作能力变化的方差分数是多少？ （2分）**

A. 48.0% B. 49.4% C. 70.3% D. 37.8%

1. **应该保留哪些变量用于简化的回归模型？ （2分）**

A. 治疗（TREAT）、年龄（AGE）、贝塔受体阻滞剂（B\_BLOCK）和吸烟（SMOKE）

B. 年龄（AGE）、贝塔受体阻滞剂（B\_BLOCK）和吸烟（SMOKE）

C. 治疗（TREAT）和年龄（AGE）

D. 治疗（TREAT）、年龄（AGE）、贝塔受体阻滞剂（B\_BLOCK）、吸烟（SMOKE）和糖尿病（DIAB）

**模型3：运行一个简化的回归模型，简化到足以适应数据。这个简化的模型应该包括主要的解释变量和试验的结果，并大致预测未来心肌梗塞患者结果变量的平均值。（问题10-19）**

1. **比较模型2和模型3，哪个模型更可取？ （2分）**

A. 之前的模型II更可取，因为那个模型的R平方更高。

B. 模型III更可取，因为那个模型的调整R平方更高，并且包含的变量更少

C. 之前的模型II更可取，因为那个模型有更多的解释变量。

D. 模型III更可取，因为那个模型的调整R平方更低。

1. **预测变量中最强的贡献者是谁？ （2分）**

A. β受体阻滞剂（B\_BLOCK）

B. 年龄（AGE）

C. 所有解释变量同样强。

D. 治疗（TREAT）

1. **患者每增加5岁，对工作能力变化的平均影响是多少？ （2分）**

A. -2.2瓦特

B. -12.3瓦特

C. -10.9瓦特

D. 我们需要更多信息来回答这个问题

1. **关于年龄效应，比较模型3和模型2（2分）**

A. 斜率系数的符号不同

B. 关于年龄的斜率系数没有相关变化

C. 斜率系数变化了-1.22单位

D. 斜率系数变化了2.18单位

**预测不同治疗的个体预期的平均WEFF**

许多医生希望预测来到他们诊所的患者的结果。基于背景因素，是否可能使用回归分析的结果来预测结果？注意：根据你的模型3回答以下问题

1. **假设我们有一个48岁的患者，吸烟但不使用贝塔受体阻滞剂。该患者接受标准治疗。这样的患者的预期WEFF是多少？ （2分）**

A. 45.2瓦特 B. 56.5瓦特 C. 48.7瓦特 D. 无法估计

1. **对于一个75岁的患者，不吸烟，使用β受体阻滞剂，接受强化治疗，预期的WEFF是多少？ （2分）**

A. -6.02瓦特 B. 3.8 C. -16.1瓦特 D. 无法估计

**评估回归模型与适当的残差图。检查线性、等残差方差和正态分布残差的假设：**

如果你使用SPSS，你必须保存原始（未标准化）残差和预测值。用残差在y轴上，预测值在x轴上制作散点图。

1. **绘制残差与预测值的图。关于方差齐的假设的结论是什么？ （2分）**

A. 由于存在扇形图案，意味着方差不等，假设没有得到满足。

B. 假设没有得到满足，因为残差有很大的变异性，一些观测的残差值很小，而其他观测的残差值要大得多。

C. 由于预测值增加时残差的分布没有系统性增加或减少，假设得到满足。 、

D. 由于所有值都围绕零值居中，假设得到满足。

1. **在Q-Q图或P-P图中绘制残差（或直方图），以检查正态分布的假设 （2分）**

A. 明显不满足正态假设

B. 这些数据的正态假设不相关

C. 正态假设大致满足

D. 残差没有分布，因为它只是随机的

1. **假设你已经包含了最重要的解释变量，模型是否有效？ （2分）**

A. 大约是的

B. 不能在这个数据上计算有效性

C. 一点也不是

D. 由于模型缺少重要变量，我们无法确定模型的有效性

1. **应该呈现哪种治疗效应的差异？ （2分）**

A. 模型III的调整效应

B. 模型II的调整效应

C. t检验会话中的未调整效应

D. 组间均值的差异，因为描述性统计总是正确的。