1. 上周听了苗老师的讲座受益匪浅，尤其整体的思路感触很深，从个人的理解进行回顾，以及一点自己的想法
2. 做causal effect需要依赖三个假设，其中有一个叫 Unconfounderness，也就是没有未观测到的混杂因素，现实中很难实现
3. 工具变量法一直是常用解决办法，但工具变量法又引入了一个新的假设Exclusion，要求Z不能直接影响Y，我们知道这种情况现实世界也是很难找到的
4. 还有另一种解决办法，也就是代理变量法，线性情况下，有四种代理变量模型，需满足不同先验识别因果效应，这里举一个例子X代表启智计划，Y代表儿童认知，U代表社会经济地位，图1(a)W是父亲的职业，可作为U的代理变量，需要极高的先验知识。图1(b)W表示父亲的职业，Z表示家庭收入。(C)(D)更符合现实情况
5. 拓展到非线性的情况，苗老师的论文里进行了各类情况的更详细描述，我们发现(d,e,f)更贴近于现实情况，(f)是最普遍版本，X和Z之间的关系由于独立性的满足是可以变换的Negative controls克服先验p(w/u)
6. 建立两个Confounding bridge function（NCO和NCE），NCO举一例急性应激对心脏病死亡率影响，心脏和外部原因引起的死亡率在增加年雅典地震后的那几天，然而，地震造成的急性压力不太可能迅速导致癌症死亡。在一项平行的分析中，他们没有发现癌症死亡率的风险的增加，这是支持没有混淆的证据，并强化了他们的主张，即急性压力会增加来自心脏病的死亡率。癌症死亡率是阴性对照结果，用于测试是否存在混杂偏倚，并评估因果关联的合理性。
7. NCE举一例，研究人员对宫内暴露对后代结果的影响相当感兴趣，例如，母亲在怀孕期间吸烟对后代出生体重、哮喘和肥胖的影响。如果存在宫内因果机制，那么母亲暴露预计会对后代结果产生影响，但以母亲暴露为条件，父亲暴露不应影响后代结果。将父亲吸烟作为阴性对照暴露，以调整母亲吸烟对后代出生体重和晚年体重指数的宫内影响。
8. Negative controls最新应用，TND测试阴性设计流程，减少混杂偏差，有向无环图说明了寻求医疗保健行为的选择偏差。V流感疫苗，I感染，HS寻求医疗保健，C混杂 A)寻求医疗保健HS混淆了V型疫苗接种和流感状态I之间的关系，也影响了对研究t的检测和选择。流感状态也可能影响寻求护理和接受检测的人。V→I关系的其他混杂因素用C表示，也可能影响HS。只有接受了流感（T = 1）检测的患者被纳入研究，导致对撞机偏倚。B)HS控制（HS = 1）阻塞了偏置路径。
9. 同样这个问题可以用Negative controls来做，检验阴性设计中变量的因果关系。(a)说明变量之间的因果关系在一般人群测试阴性设计，并使用(b)说明假设隐含的常见方法估计兽医的研究对象有相同的求医行为（HSB）(c)表明，如果HSB仍然部分未被观察到，那么后门路径A←HSB→Y和A←HSB→S = 1←Y分别表示未测量的混杂偏差和选择偏差。其他未测量的混杂因素，如医疗保健工作者的职业和以前的感染，在a和Y之间打开了额外的后门路径，并导致额外的混杂偏差。(d)显示了(c)的简化DAG，它将未测量的混杂因素合并为一个单一变量U。(e)说明了我们利用阴性控制暴露Z和结果w来估计VE的方法。虚线箭头表示不需要的影响。(f)显示了一个使用a→S箭头的场景，在其他假设下仍然可以将因果比确定。
10. 特异性问题，1950实验对比发现，吸烟与肺癌的关系比管道吸烟更密切。没有发现与吸入有明显的关联，2005特异性的标准要求一个原因导致单一效应，而不是多重效应。这一论点经常被提出来反驳似乎与无数影响有关的暴露的因果解释——例如，那些试图证明吸烟是导致肺癌的原因的人。不幸的是，该标准作为一般规则是无效的。所以吸烟和肺癌到底是不是因果关系呢？
11. P-value方法进行验证，值越小显著性越强，式子中我们发现如果X是u的无偏估计，没问题，如果X有偏倚，则必须从其他数据中估计偏差，并在进行测试之前从X中删除X先进行去偏。也就说是P是关于抽样误差的，而不是关于偏差的。
12. 苗老师还没有最新相关论文发表，只能参考ppt，个人理解讲。k个X，p个Y，除了所选X中的其他X
13. 线性模型下的关系