Question 2:

1. Let

$$e_t(x) = E[\mathbb{I}\{T = t\}|X = x] = P(T = t|X = x)$$
$$q_t(x) = \frac{\mathbb{I}(T = t)}{e_t(x)} = \frac{\mathbb{I}(T = t)}{E[\mathbb{I}\{T = t\}|X = x]}$$

We assume Consistency: $Y = TY_1 + (1 - T)Y_0$ so for T=t we get that $\mathbb{I}\{T = t\} \cdot Y = \mathbb{I}\{T = t\} \cdot Y_t$ (*)

Ignorability: $(Y_0, Y_1) \perp T|X$ (**)

$$E[q_t(x)Y] = E\left[\frac{\mathbb{I}\{T = t\}}{e_t(x)}Y\right] =^{(*)} E\left[\frac{\mathbb{I}\{T = t\}}{e_t(x)}Y_t\right] =^{(Law\ of\ total\ expectation)}$$
$$= E_{x\sim p(x)}\left[E\left[\frac{\mathbb{I}\{T = t\}}{e_t(x)}Y_t|X\right]\right].$$

The inner expression:

$$E\left[\frac{\mathbb{I}\{T = t\}}{e_t(x)}Y_t|X\right] =^{e_t(x)\ is\ const} \frac{1}{e_t(x)}E[\mathbb{I}\{T = t\} \cdot Y_t|X] =^{(**)} \frac{1}{e_t(x)} \cdot E[\mathbb{I}\{T = t\}|X]$$
$$\cdot E[Y_t|X] =^{(e_t(x)\ defenition)}= \frac{1}{e_t(x)} \cdot e_t(x) \cdot E[Y_t|X] = E[Y_t|X].$$

We plug back into the original expression and we get:

$$E[q_t(x)Y] = E_{x\sim p(x)}\left[E\left[\frac{\mathbb{I}\{T = t\}}{e_t(x)}Y_t|X\right]\right] = E_{x\sim p(x)}\big[E[Y_t|X]\big] = E[Y_t].$$

2. We will denote the covariate vector for a pregnant woman $X = x_{preg}$ and we know that in the trial there were no pregnant women, so: $Pr\left(T = 1\middle|X = x_{preg}\right) = 0 \Rightarrow e_1\left(x_{preg}\right) = 0.$ This means $q_1\left(x_{preg}\right)$ is undefined (division by zero), and the same for $q_0\left(x_{preg}\right)$.

   That information violates the positivity of the trial, and we cannot say anything about the causal effect that the drug has on pregnant women.

Question 3:

1) The SUTVA assumption does **NOT** hold.
For example, in treatments Y00000010 and Y00000011, the only change is that 'Steph Curry' got a basketball, but it affects 'Seth Curry', who got 0 in the first treatment and 1 with the second one.

2) Because each family in our data has 2 kids, and we assume the SUTVA holds between the different families, we can define the treatment as the number of basketballs each family got.
For a family with kids i,j we can write the new treatment as $T_i + T_j \in \{0,1,2\}$.
And the potential outcomes, that is, the number of children who got a full scholarship in each family - $Y_f = Y_i + Y_j \in \{0,1,2\}$.

3) First, we will create an updated table:

| | $Y_f$(0) | $Y_f$ (1) | $Y_f$ (2) |
|---|---|---|---|
| Ball | 0 | 1 | 0 |
| Plumlee | 0 | 2 | 0 |
| Lopez | 0 | 1 | 0 |
| Curry | 0 | 0 | 2 |

$$ATE(1-0) = \frac{1}{4}\sum_f [Y_f(1) - Y_f(0)] = \frac{1+2+1+0}{4} = 1$$

$$ATE(2-0) = \frac{1}{4}\sum_f [Y_f(2) - Y_f(0)] = \frac{0+0+0+2}{4} = \frac{1}{2}$$

$$ATE(2-1) = \frac{1}{4}\sum_f [Y_f(2) - Y_f(1)] = \frac{-1-2-1+2}{4} = -\frac{1}{2}$$