A.

From two conditions $P(y_i{=}1|s_i{=}0){=}P(y_i{=}1)$ and $P(\mathbf{x}|s_i{=}1){=}P(\mathbf{x}|y_i{=}1)$,

$$\begin{aligned}
&P(f_i{<}f_j \mid s_i{=}1,s_j{=}0)\\
&=P(f_i{<}f_j \mid y_i{=}1,s_j{=}0)\\
&=P(y_j{=}0 \mid y_i{=}1,s_j{=}0)P(f_i{<}f_j \mid y_i{=}1,y_j{=}0,s_j{=}0)\\
&+P(y_j{=}1 \mid y_i{=}1,s_j{=}0)P(f_i{<}f_j \mid y_i{=}1,y_j{=}1,s_j{=}0)\\
&=P(y_j{=}0 \mid y_i{=}1)P(f_i{<}f_j \mid y_i{=}1,y_j{=}0)\\
&+P(y_j{=}1 \mid y_i{=}1)P(f_i{<}f_j \mid y_i{=}1,y_j{=}1). \qquad (20)
\end{aligned}$$

Similarly,

$$\begin{aligned}
&\frac{1}{2}P(f_i{=}f_j \mid s_i{=}1,s_j{=}0)\\
&=\frac{1}{2}P(y_j{=}0 \mid y_i{=}1)P(f_i{=}f_j \mid y_i{=}1,y_j{=}0)\\
&+\frac{1}{2}P(y_j{=}1 \mid y_i{=}1)P(f_i{=}f_j \mid y_i{=}1,y_j{=}1). \qquad (21)
\end{aligned}$$

Combining them, we obtain

$$\begin{aligned}
&R_X(i,j)\\
&=P(f_i{<}f_j \mid s_i{=}1,s_j{=}0)+\frac{1}{2}P(f_i{=}f_j \mid s_i{=}1,s_j{=}0)\\
&=(1-\pi_{ij})R(i,j)+\pi_{ij}R_{-X}(i,j). \qquad (22)
\end{aligned}$$

B.

From (13),

$$\begin{aligned}
&P(y_i{=}1,y_j{=}0)R(i,j)\\
&=\frac{P(y_i{=}1,y_j{=}0)}{P(y_j{=}0|y_i{=}1)}\left\{R_X(i,j)-P(y_j{=}1|y_i{=}1)R_{-X}(i,j)\right\}\\
&=P(y_i{=}1)R_X(i,j)-P(y_i{=}1,y_j{=}1)R_{-X}(i,j), \qquad (23)
\end{aligned}$$

Plugging the equation above and (14) into (6), we obtain

$$\begin{aligned}
&\mathcal{L}_{\text{rank}}\\
&=\sum_{1\le i<j\le m} P(y_i{=}1,y_j{=}0)R(i,j)+P(y_i{=}0,y_j{=}1)R(j,i)\\
&=\sum_{1\le i<j\le m} P(y_i{=}1)R_X(i,j)+P(y_j{=}1)R_X(j,i)\\
&-P(y_i{=}1,y_j{=}1)\left\{R_{-X}(i,j)+R_{-X}(j,i)\right\}\\
&=\sum_{1\le i<j\le m} P(y_i{=}1)R_X(i,j)+P(y_j{=}1)R_X(j,i)\\
&-P(y_i{=}1,y_j{=}1), \qquad (24)
\end{aligned}$$

C.

Because

$$\begin{aligned}
&P(y_i{=}1)-P(s_i{=}1,s_j{=}0)\\
&=P(y_i{=}1,s_i{=}0)+P(y_i{=}1,s_i{=}1)-P(y_i{=}1,s_i{=}1,s_j{=}0)\\
&=P(y_i{=}1,s_i{=}0)+P(s_i{=}1,s_j{=}1), \qquad (25)
\end{aligned}$$

from (15) and (9), we obtain

$$\begin{aligned}
&\mathcal{L}_{\text{rank}}-\hat{\mathcal{L}}_{\text{rank}}\\
&=\sum_{1\le i<j\le m} P(y_i{=}1)R_X(i,j)+P(y_j{=}1)R_X(j,i)-\text{const}\\
&-\sum_{1\le i<j\le m} P(s_i{=}1,s_j{=}0)R_X(i,j)+P(s_i{=}0,s_j{=}1)R_X(j,i)\\
&=\sum_{1\le i<j\le m} P(s_i{=}1,s_j{=}1)\left\{R_X(i,j)+R_X(j,i)\right\}\\
&+P(y_i{=}1,s_i{=}0)R_X(i,j)+P(y_j{=}1,s_j{=}0)R_X(j,i)\\
&+\text{const}. \qquad (26)
\end{aligned}$$

The first term is proportional to the ratio of samples having both i-th and j-th label. Second and third terms are proportional to ratio of samples, which is not labeled even if they are positive.

D.

When labels are given completely, the loss function which should be minimized is (6),

$$\begin{aligned}
&\mathcal{L}_{\text{rank}}\\
&=\sum_{1\le i<j\le m} P(y_i{=}1,y_j{=}0)R(i,j)+P(y_i{=}0,y_j{=}1)R(j,i). \qquad (27)
\end{aligned}$$

However, the loss function derived based on "case-controlled" assumption is (15)

$$\begin{aligned}
&\mathcal{L}_{\text{rank}}\\
&=\sum P(y_i{=}1)R_X(i,j)+P(y_j{=}1)R_X(j,i)-P(y_i{=}1,y_j{=}1) \qquad (28)
\end{aligned}$$

Considering the case in which label deficit does not exist, because $R(i,j){=}R_X(i,j)$,

$$\begin{aligned}
&\mathcal{L}_{\text{rank-false}}\\
&=\sum P(y_i{=}1)R(i,j)+P(y_j{=}1)R(j,i)-P(y_i{=}1,y_j{=}1). \qquad (29)
\end{aligned}$$

Their mutual difference is

$$\begin{aligned}
&\mathcal{L}_{\text{rank-false}}-\mathcal{L}_{\text{rank}}\\
&=\sum P(y_i{=}1,y_j{=}1)\left\{R(i,j)+R(j,i)-1\right\}. \qquad (30)
\end{aligned}$$

This error is proportional to the joint probability that both i-th class and j-th class are positive. However, this error is cancelled if we use symmetric surrogate loss.