Certainly! Let's go through the entire calculation using the **original scenario** where:

- **Class 0** = Disease (Positive Class)
- **Class 1** = Healthy (Negative Class)

Confusion Matrix:

	Predicted Class 0 (Disease)	Predicted Class 1 (Healthy)
Actual Class 0 (Disease)	20	30
Actual Class 1 (Healthy)	10	940

Step-by-Step Calculation of Metrics:

1. True Positives (TP):

- These are the cases where the model correctly predicts **Disease** (Class 0) as **Disease** (Class 0).
- From the confusion matrix: **TP = 20**.

2. False Positives (FP):

- These are the cases where the model incorrectly predicts **Healthy** (Class 1) as **Disease** (Class 0).
- From the confusion matrix: **FP = 10**.

3. False Negatives (FN):

- These are the cases where the model incorrectly predicts **Disease** (Class 0) as **Healthy** (Class 1).
- From the confusion matrix: **FN = 30**.

4. True Negatives (TN):

- These are the cases where the model correctly predicts **Healthy** (Class 1) as **Healthy** (Class 1).
- From the confusion matrix: **TN = 940**.

Now that we have all the necessary values, let's calculate the key metrics: **Precision**, **Recall**, and **F1-Score**.

Precision:

Precision tells us how many of the predicted **positive** cases (Disease) are actually correct. It is calculated as:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} = \frac{20}{20 + 10} = \frac{20}{30} = 0.67$$

So, **Precision** is **0.67**.

Recall (Sensitivity or True Positive Rate):

Recall tells us how many of the **actual positive** cases (Disease) were correctly identified by the model. It is calculated as:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} = \frac{20}{20 + 30} = \frac{20}{50} = 0.40$$

So, **Recall** is **0.40**.

F1-Score:

The F1-Score is the harmonic mean of **Precision** and **Recall**, and it gives us a balanced measure of the model's performance by considering both false positives and false negatives. It is calculated as:

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} = 2 \times \frac{0.67 \times 0.40}{0.67 + 0.40} = \frac{0.536}{1.07} = 0.50$$

So, **F1-Score** is **0.50**.

Accuracy:

Accuracy tells us the proportion of correctly predicted cases (both positive and negative) out of all cases. It is calculated as:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} = \frac{20 + 940}{20 + 940 + 10 + 30} = \frac{960}{1000} = 0.96$$

So, **Accuracy** is **0.96**.

Summary of Metrics:

Metric	Value
Precision	0.67
Recall	0.40
F1-Score	0.50
Accuracy	0.96

Conclusion:

- **Precision** (0.67) indicates that out of all the instances predicted as **Disease**, 67% were actually **Disease**.
- **Recall** (0.40) indicates that the model correctly identified only 40% of the actual **Disease** cases, meaning it missed 60% of the **Disease** cases (these are **False Negatives**).
- **F1-Score** (0.50) is a balanced measure that considers both **Precision** and **Recall**, showing that the model has a moderate performance in identifying **Disease** cases.
- **Accuracy** (0.96) is very high, but this can be misleading in an imbalanced dataset because the model could be predicting **Healthy** (Class 1) correctly most of the time, but it fails to identify enough **Disease** (Class 0) cases.

Why False Negatives (FN) Matter:

In medical diagnosis, **False Negatives (FN)** are critical because failing to detect a **Disease** case (Class 0) means the patient will not receive the necessary treatment. This could have severe consequences for the patient's health. Even though the **Accuracy** is high (96%), the model misses **30** Disease cases (FN), which should be a major concern in healthcare applications.