# Email Marketing Content A/B test

# Test specifications

- Target Audience- emailable costumers (users that have not opted out of marketing email)

- Groups- two of identical size (214,206 costumers each):
  o Test: received an e-mail containing content 2
  o Control : received an e-mail containing content 1

- 4 period were evaluated in order to score the test:

| Period name | Dates | Comments |
|---|---|---|
| **Pre-Campaign** | 26/12/2017 – 8/1/2018 | significant differences in this period might imply that the groups were not divided well, as there are other bias that might affect the test results other than the test itself. |
| **Campaign** | 9/1/2018 – 13/1/2018 | |
| **Post** | 14/1/2018 – 27/1/2018 | Significant difference might imply of demand being pulled forward |
| **Campaign through post combined** | 9/1/2018- – 27/1/2018 | includes action made due to the test campaign in campaign and post periods |

# Method – T test

- **Response metric** (for conversions):

  **Response metric, t statistic** $= [\text{SQROOT}(\text{mean1-mean2})^2]/[(\text{std dev1})^2 + (\text{std dev2})^2]$

- **Continuous metric** (for averages):

  **Continuous metric, t statistic** $= A/(B*C)$

  $A = \text{ABSOLUTE}[(\text{mean1-mean2})$

  $B = \text{SQROOT}\{[(\text{aud size1})*(\text{variance1}) + (\text{aud size2})*(\text{variance2})] / (\text{aud size1} + \text{aud size2} -2)\}$

  $C = \text{SQROOT}[(\text{Aud size1} + \text{Aud size2})/(\text{Aud size1})*(\text{Aud size2})]$

# Metrics

**Success metrics:**
- **Conversions\*-**
  - **click through rate-** of unique openers
  - **complains rate-** of unique openers

- **Averages-**
  - revenue per buyer
  - revenue per order
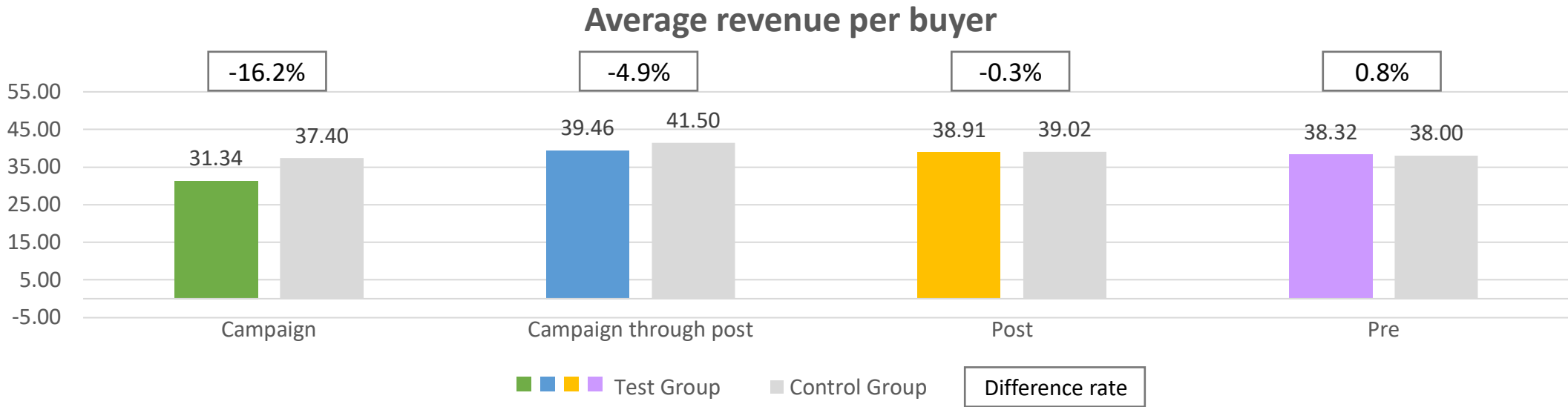
**Validation metrics\*\*:**
- **opening percentage-** as being expose to the test subject (the email content) demands opening it first, the opening rates are irrelevant for measuring the test success. However**,** a significant difference between the groups might suggest that the groups were not devided properly and there are bias.
- **click through rate of sends-** for the same reason as the opening rate, this metric should be compared to click through rate of unique openers. significant difference in both might imply that the groups are bias
- **Complains rate of sends-** for the same reason as the opening rate, this metric should be compared to Complains rate of unique openers. significant difference in both might imply that the groups are bias.

\*   The conversions were measured only for the entire time starting campaign launch date (Campaign through post combined)
\*\*   Validation metrics are calculated in order to ensure that the test groups has no bias that might effect the results
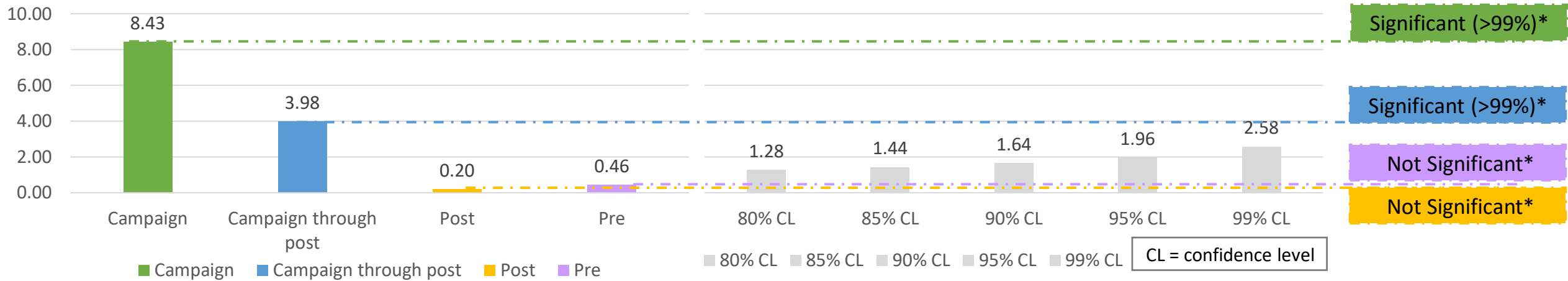
# Success metrics: Average revenue per buyer –
the test group was significantly lower in the **campaign** and **Campaign through post** periods (confidence of above 99%) there's no sign of groups bias in the **pre** period, or demand being pulled forward to the **post** period



**Average revenue per buyer**

| | Campaign | Campaign through post | Post | Pre |
|---|---|---|---|---|
| Difference rate | -16.2% | -4.9% | -0.3% | 0.8% |
| Test Group | 31.34 | 39.46 | 38.91 | 38.32 |
| Control Group | 37.40 | 41.50 | 39.02 | 38.00 |

**Average revenue per buyes – t-statistic**

t-statistic values: Campaign 8.43, Campaign through post 3.98, Post 0.20, Pre 0.46

**t-critical**

80% CL 1.28, 85% CL 1.44, 90% CL 1.64, 95% CL 1.96, 99% CL 2.58

Significant (>99%)*
Significant (>99%)*
Not Significant*
Not Significant*

CL = confidence level

\* Significance requires: t statistic > t critical

# Success metrics: Average revenue per order –
## similar results for revenue per order data

### Average revenue per order



| -16.1% | -6.4% | -1% | 0% |

Campaign: 27.80 (Test Group), 33.14 (Control Group)
Campaign through post: 30.04 (Test Group), 32.09 (Control Group)
Post: 31.37 (Test Group), 31.69 (Control Group)
Pre: 30.91 (Test Group), 30.91 (Control Group)

Test Group | Control Group | Difference rate

### Average revenue per order – t-statistic          t-critical



Campaign: 7.90 — Significant (>99%)*
Campaign through post: 4.60 — Significant (>99%)*
Post: 0.59 — Not Significant*
Pre: 0.00 — Not Significant*

80% CL: 1.28
85% CL: 1.44
90% CL: 1.64
95% CL: 1.96
99% CL: 2.58

Campaign | Campaign through post | Post | Pre

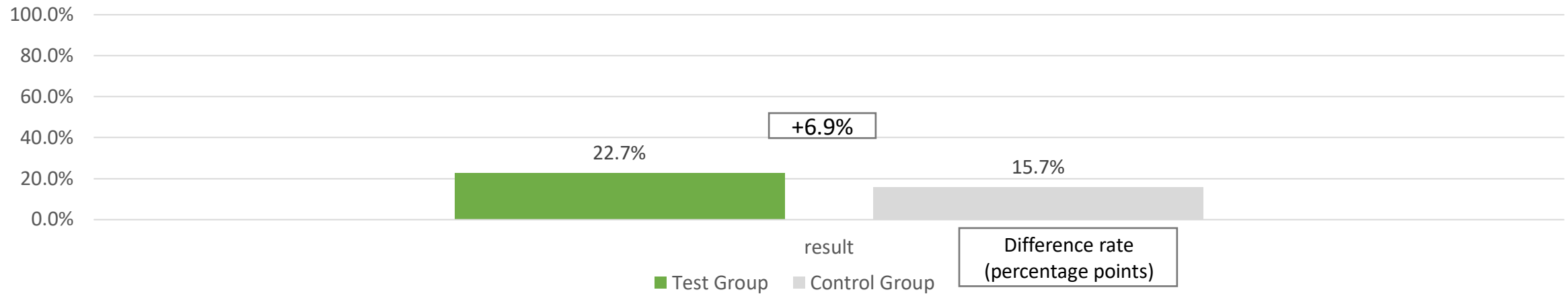80% CL | 85% CL | 90% CL | 95% CL | 99% CL | CL = confidence level

* Significance requires: t statistic > t critical

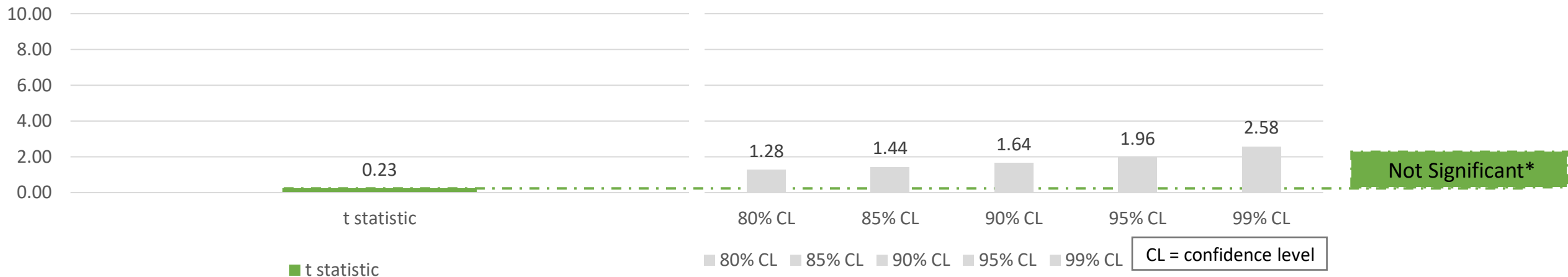# Success metrics: click through rate of unique openers conversion –
There is no significant differences between the group, concluding that the test had no significant impact on the value

## click through rate of unique openers



100.0%
80.0%
60.0%
40.0%
20.0%
0.0%

+6.9%

22.7%

15.7%

result

Difference rate
(percentage points)

■ Test Group    ■ Control Group

## click through rate of unique openers –

## t-critical



10.00
8.00
6.00
4.00
2.00
0.00

0.23

1.28    1.44    1.64    1.96    2.58

Not Significant*

t statistic

80% CL    85% CL    90% CL    95% CL    99% CL

■ t statistic

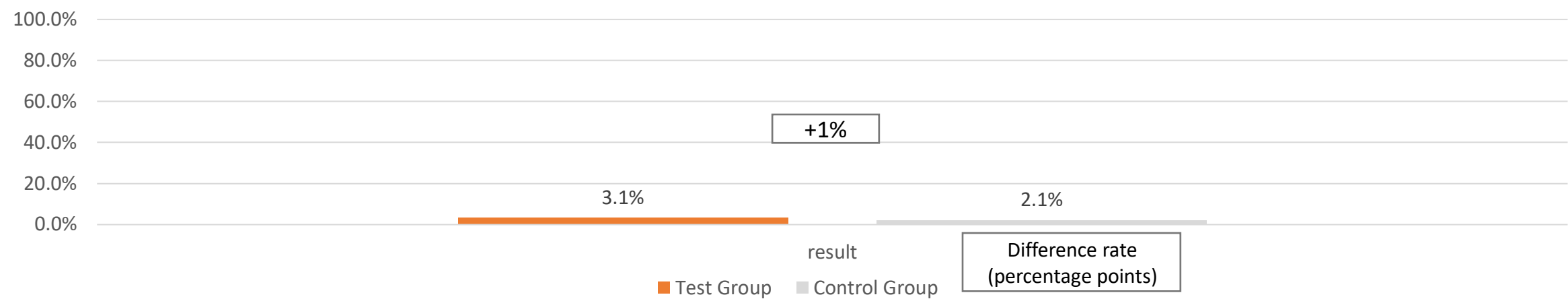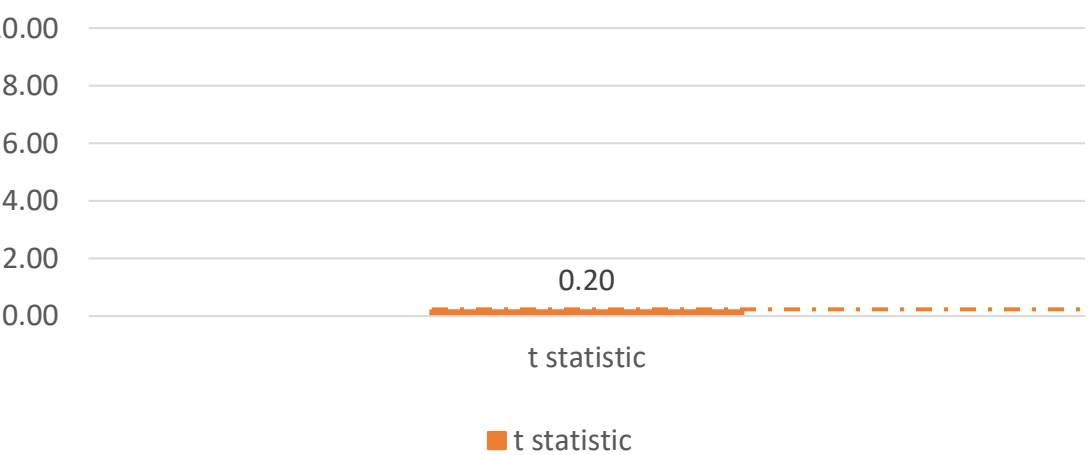■ 80% CL  ■ 85% CL  ■ 90% CL  ■ 95% CL  ■ 99% CL

CL = confidence level

* Significance requires: t statistic > t critical

# Validation metrics: click through rate of sends conversion–
## the check resulted in no significant differences as well, which implies for no group bias
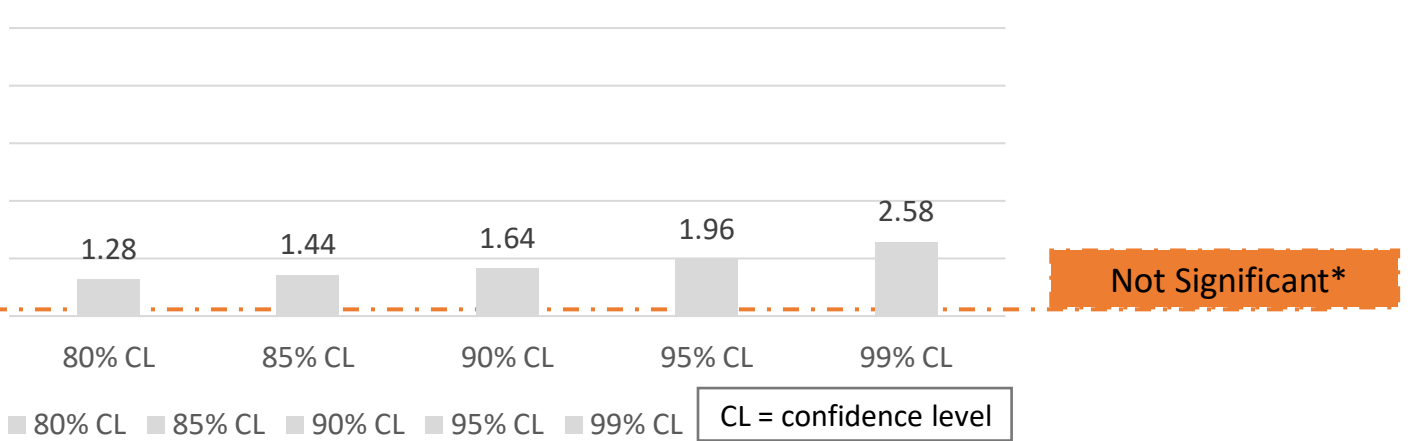
**click through rate of sends**

100.0%
80.0%
60.0%
40.0%
20.0%
0.0%

+1%

3.1%

2.1%

result

Difference rate
(percentage points)

■ Test Group   ■ Control Group

**click through rate of sends– t-statistic**

**t-critical**

10.00
8.00
6.00
4.00
2.00
0.00

0.20

1.28   1.44   1.64   1.96   2.58

Not Significant*

t statistic

80% CL   85% CL   90% CL   95% CL   99% CL

■ t statistic

■ 80% CL  ■ 85% CL  ■ 90% CL  ■ 95% CL  ■ 99% CL
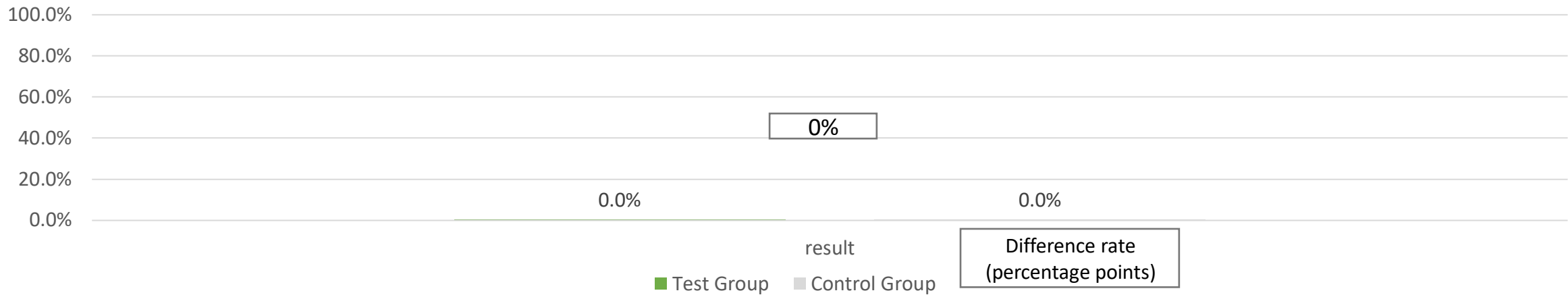
CL = confidence level

* Significance requires: t statistic > t critical

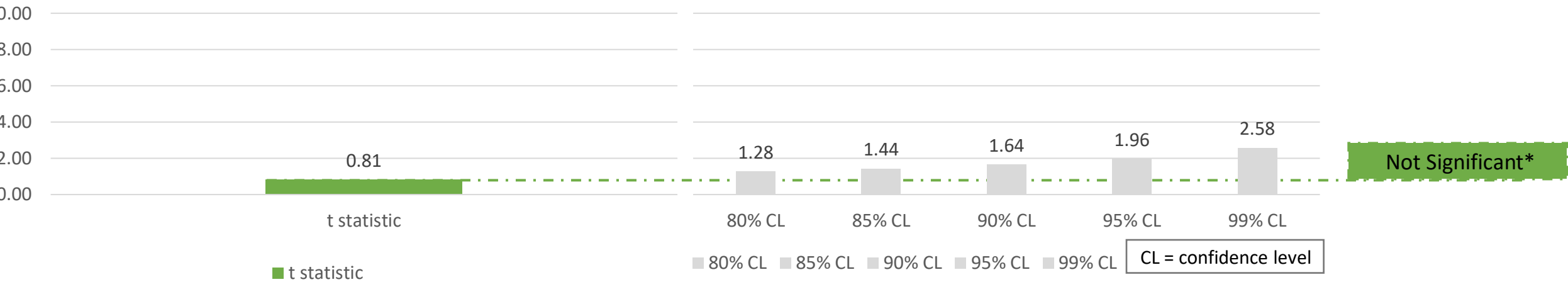# Success metrics: complains rate of unique openers conversion –
There is no significant differences between the group, concluding that the test had no significant impact on the value

## complains rate of unique openers

0%

0.0%                              0.0%

result

■ Test Group  ■ Control Group

Difference rate
(percentage points)

## complains rate of unique openers – t-statistic

### t-critical

| | | | | | |
|---|---|---|---|---|---|
| 10.00 | | | | | |
| 8.00 | | | | | |
| 6.00 | | | | | |
| 4.00 | | | | | 2.58 |
| | | 1.28 | 1.44 | 1.64 | 1.96 |
| 2.00 | 0.81 | | | | |
| 0.00 | | | | | |
| | t statistic | 80% CL | 85% CL | 90% CL | 95% CL | 99% CL |

Not Significant*

■ t statistic

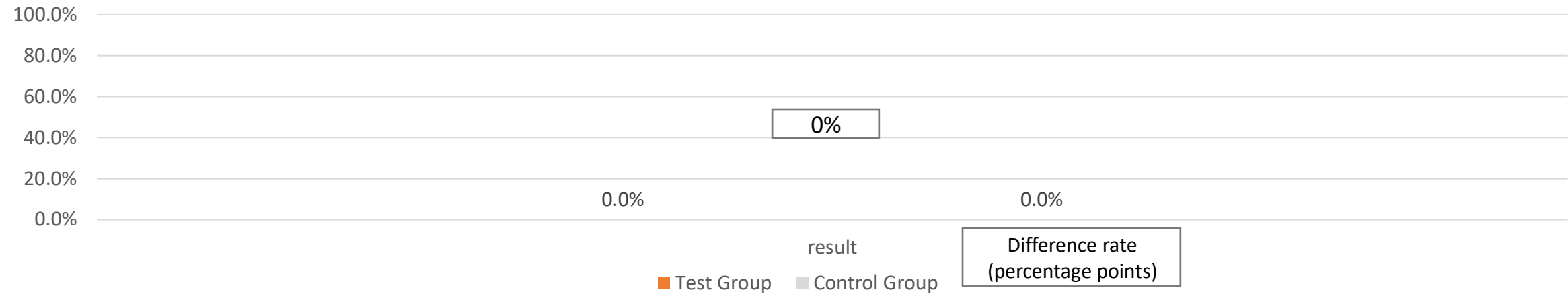■ 80% CL  ■ 85% CL  ■ 90% CL  ■ 95% CL  ■ 99% CL
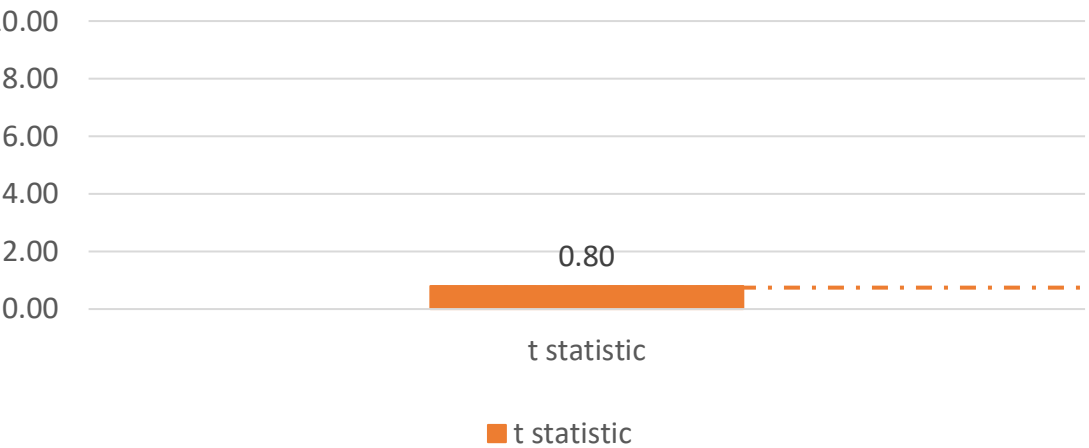
CL = confidence level

* Significance requires: t statistic > t critical

# Validation metrics: complains rate of sends conversion–
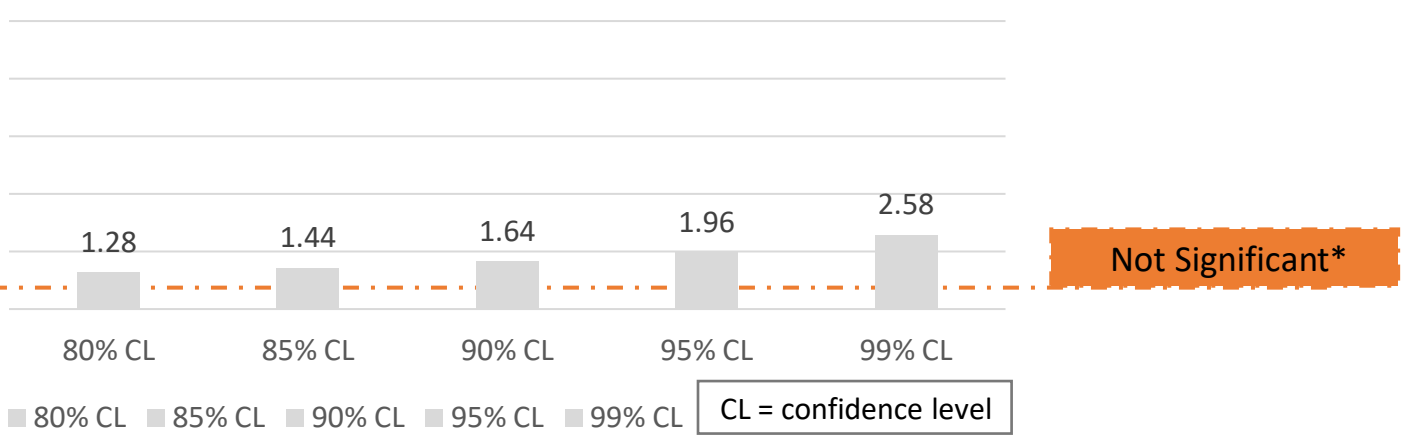## the check resulted in no significant differences as well, which implies for no group bias

**complains rate of sends**

100.0%
80.0%
60.0%
40.0%
20.0%
0.0%

0%

0.0%                                    0.0%

result

■ Test Group   ■ Control Group

Difference rate
(percentage points)

**complains rate of sends – t-statistic**                    **t-critical**

10.00
8.00
6.00
4.00
2.00
0.00

0.80

1.28      1.44      1.64      1.96      2.58

Not Significant*

t statistic

80% CL    85% CL    90% CL    95% CL    99% CL

■ t statistic

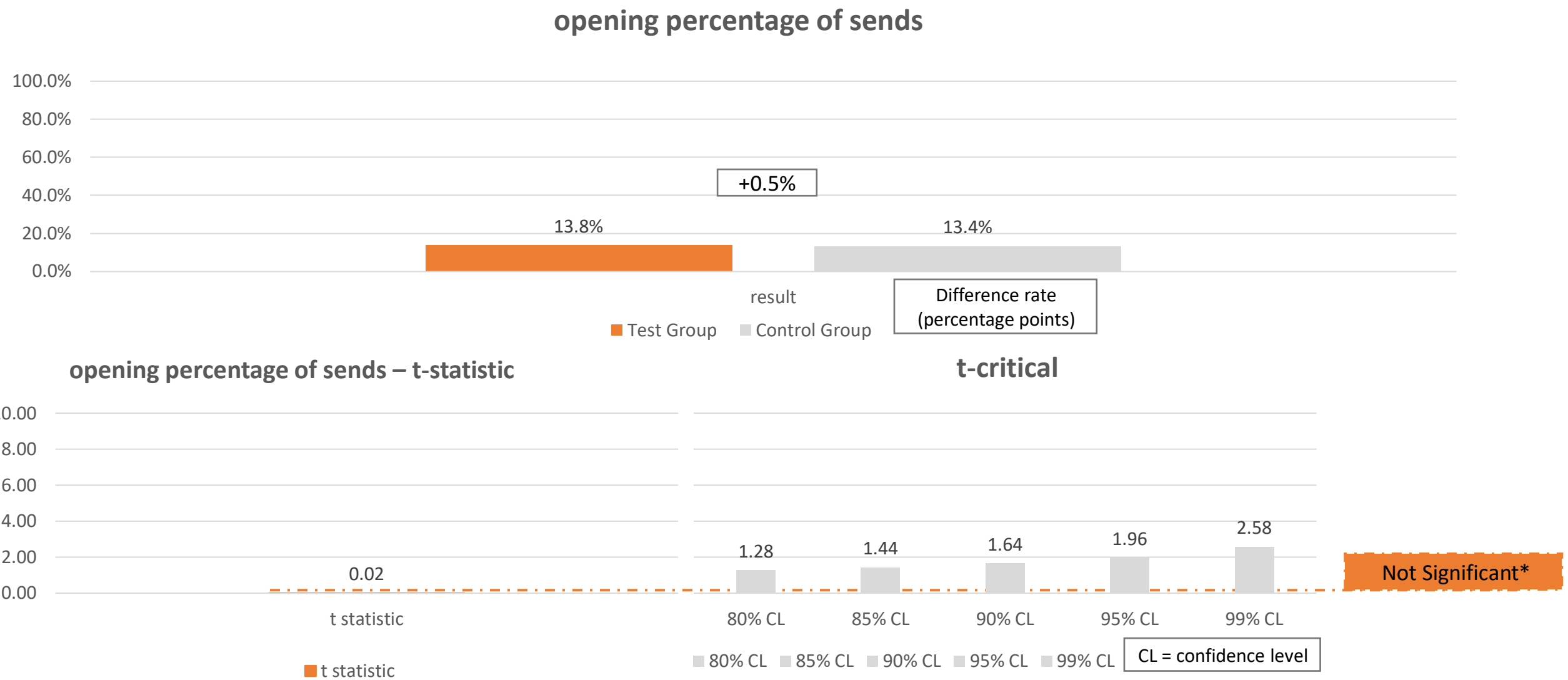■ 80% CL  ■ 85% CL  ■ 90% CL  ■ 95% CL  ■ 99% CL

CL = confidence level

* Significance requires: t statistic > t critical

# Validation metrics: opening percentage of sends conversion–
there was no significant differences found, implying for no group bias too

## opening percentage of sends

+0.5%

13.8%

13.4%

result

Difference rate
(percentage points)

■ Test Group    ■ Control Group

## opening percentage of sends – t-statistic

## t-critical

0.02

1.28    1.44    1.64    1.96    2.58

Not Significant*

t statistic

80% CL    85% CL    90% CL    95% CL    99% CL

■ t statistic

■ 80% CL  ■ 85% CL  ■ 90% CL  ■ 95% CL  ■ 99% CL

CL = confidence level

* Significance requires: t statistic > t critical

# conclusions

- There is more than 99% confidence level that Content 1 (control) is better than Content 2 (test) in terms of:
  - **Average revenue per buyer**
  - **Average revenue per order**

- There was no evidence for a preferred content in terms of:
  - **click through rate-** of unique openers
  - **complains rate-** of unique openers