R Package Risk Assessment at Novartis: Lessons Learned

PHUSE SDE
10th February 2025

NOVARTIS | Reimagining Medicine

Introduction

Principal Statistical Programmer at Novartis

Member of the Technology & Scientific Computing group

Working with the Scientific Open Source (SOS) team at Novartis

- Manage and execute the risk assessment process for R validation

Overview

Risk assessment management

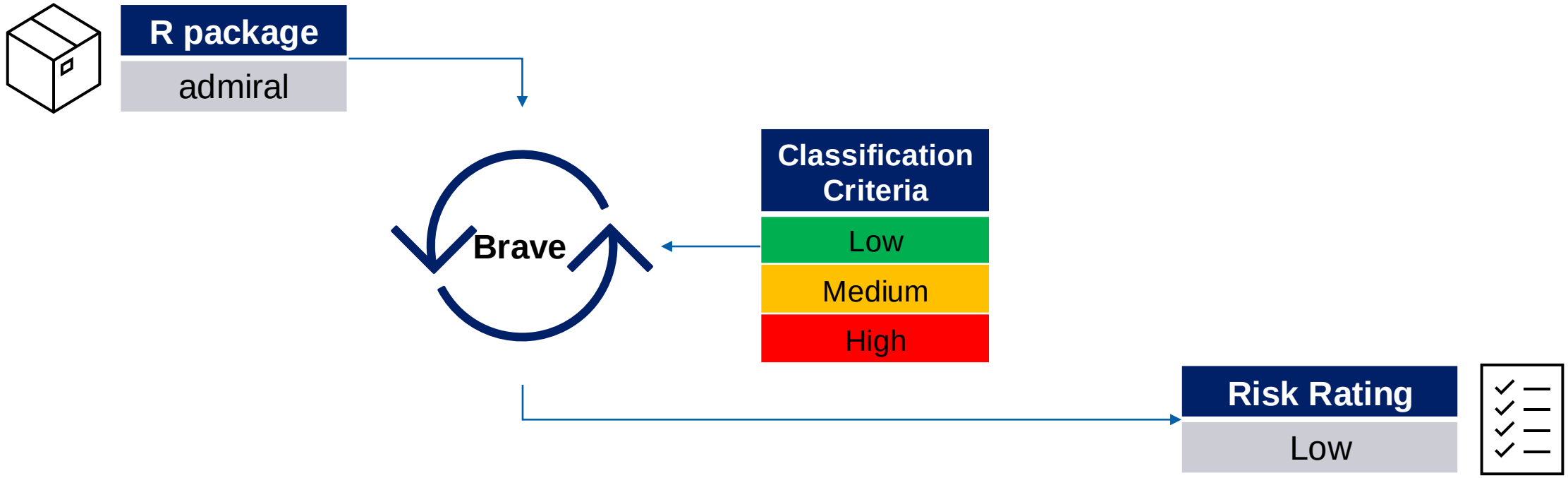
Risk assessments across releases

Causes and impact of these changes to risk assessment

Lessons learned & resulting actions our team is taking

Risk Assessment Management at Novartis

- Classify the risk level of each package as **low**, **medium**, or **high**
- Utilise internally developed R package Brave



Risk Assessment Management at Novartis

- **Low**
  - Part of base/recommends/tidyverse
  - From recognizable source or a reputable institution
  - Min. 50,000 downloads per month
  - Min. 5 reverse dependencies

- **Medium**
  - Min. 1000 downloads per month
  - Peer reviewed publications referencing package in reputable journals
  - Unit testing covering > 60%

- **High**
  - Documentation (vignettes/manuals/website/newsfeed)
  - First CRAN/Bioconductor production release of package > 1 year ago

Classification Criteria
Low
Medium
High

Risk Assessments Across Releases

Objective:

Investigate the **impact** on risk assessments across 2 releases

Risk Assessments Across Releases

Objective:

Investigate the **impact** on risk assessments across 2 releases

1. Are the risk assessments consistent across releases?

Risk Assessments Across Releases

Objective:

Investigate the **impact** on risk assessments across 2 releases

1. Are the risk assessments consistent across releases?

2. Are any changes in risk assessments impactful?

Risk Assessments Across Releases

Objective:

Investigate the **impact** on risk assessments across 2 releases

1. Are the risk assessments consistent across releases?
2. Are any changes in risk assessments impactful?
3. What happens if the version of a package changes?

Risk Assessments Across Releases: Setting Up the Analysis

Step 1

Create a master package list

Package
admiral
shinytest2
PowerTOST
plyr
...

Risk Assessments Across Releases: Setting Up the Analysis

Step 1

Create a master package list

Package
admiral
shinytest2
PowerTOST
plyr
...

Step 2

Analyse change in risk assessment for this master list of packages across two releases

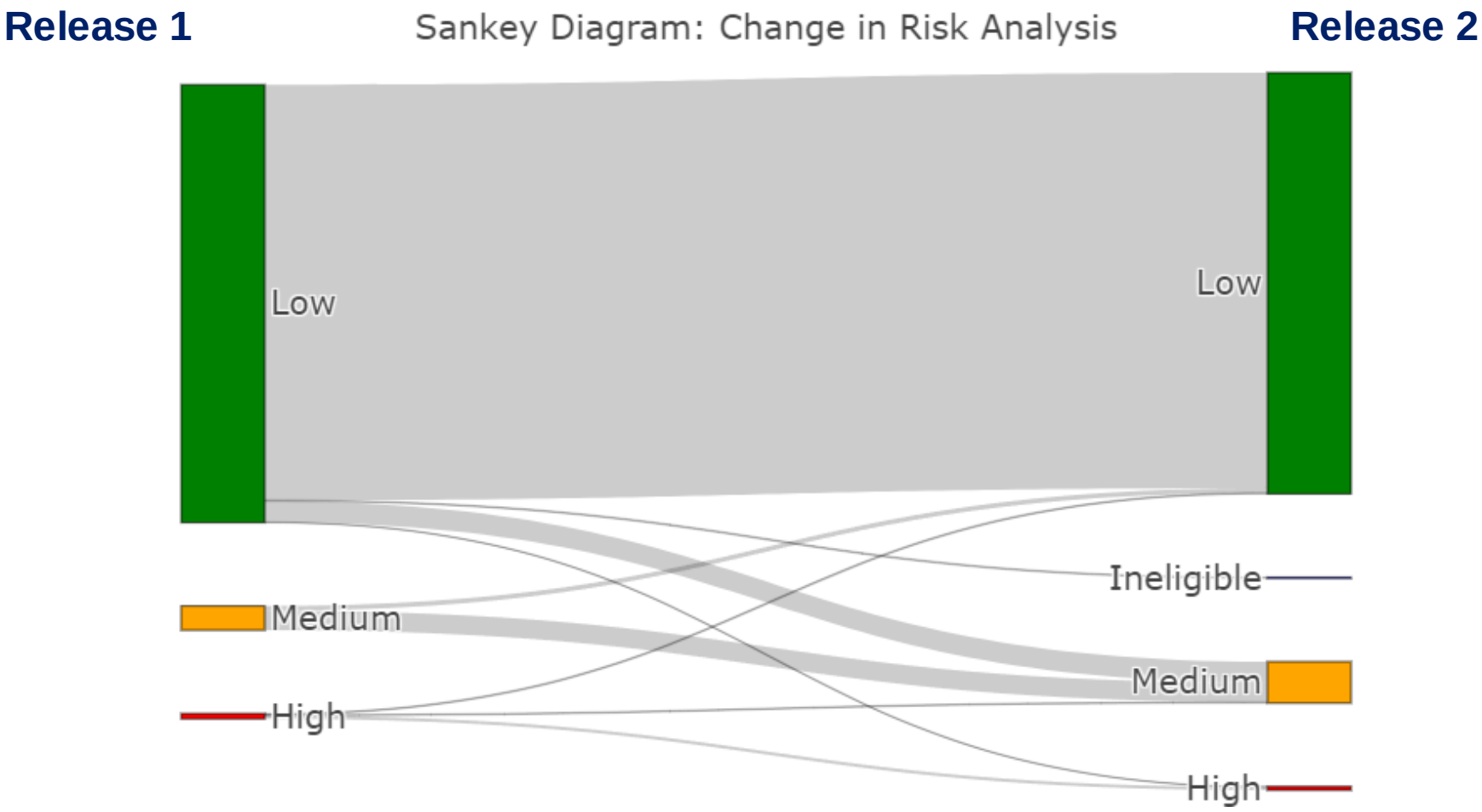
RELEASE 1

Package	Risk Assessment
admiral	Low
shinytest2	Low
PowerTOST	Low
plyr	Low
...	...

RELEASE 2

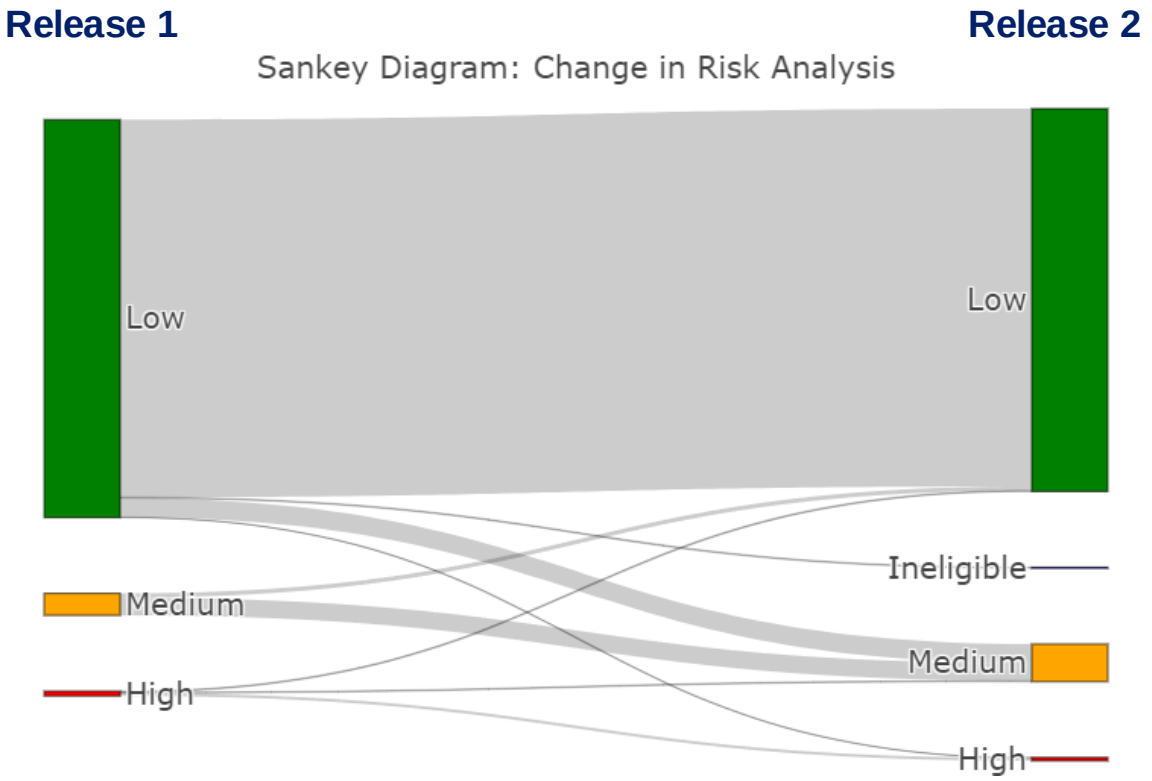
Package	Risk Assessment
admiral	Low
shinytest2	Medium
PowerTOST	High
plyr	Low
...	...

Risk Assessments Across Releases: Impact



Risk Assessments Across Releases: Impact

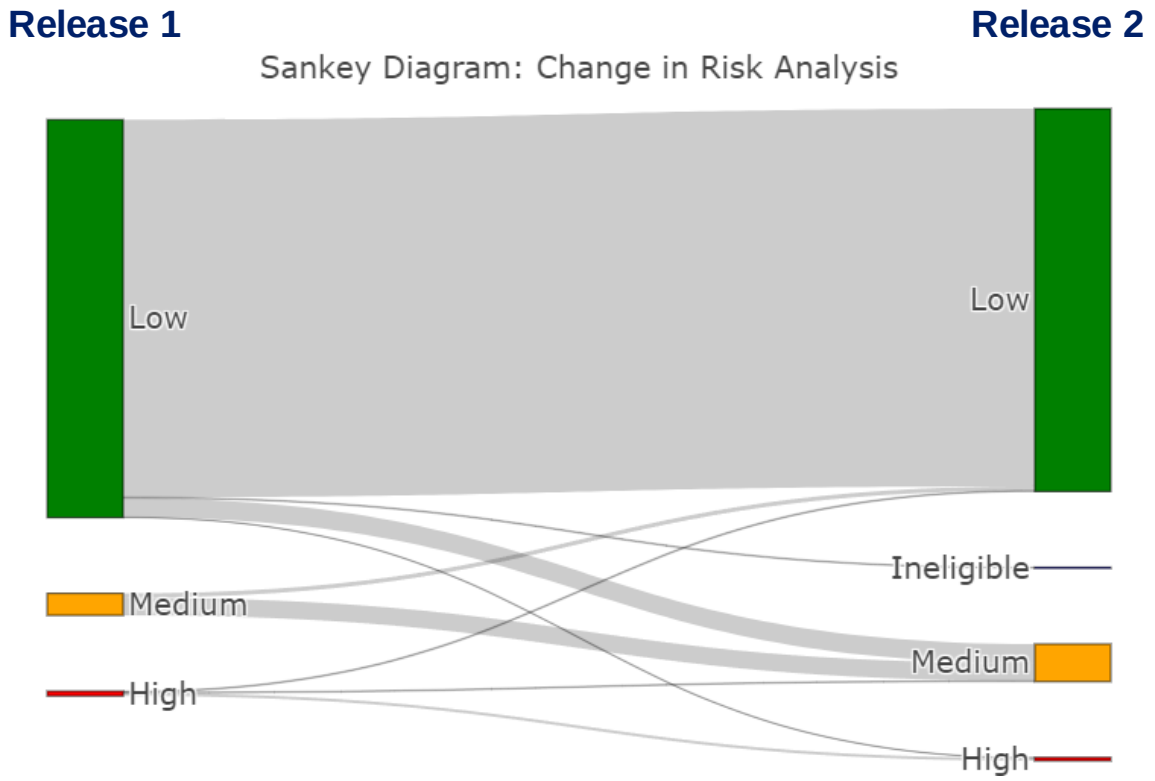
Total number of packages: 892



Risk Assessments Across Releases: Impact

Total number of packages: 892

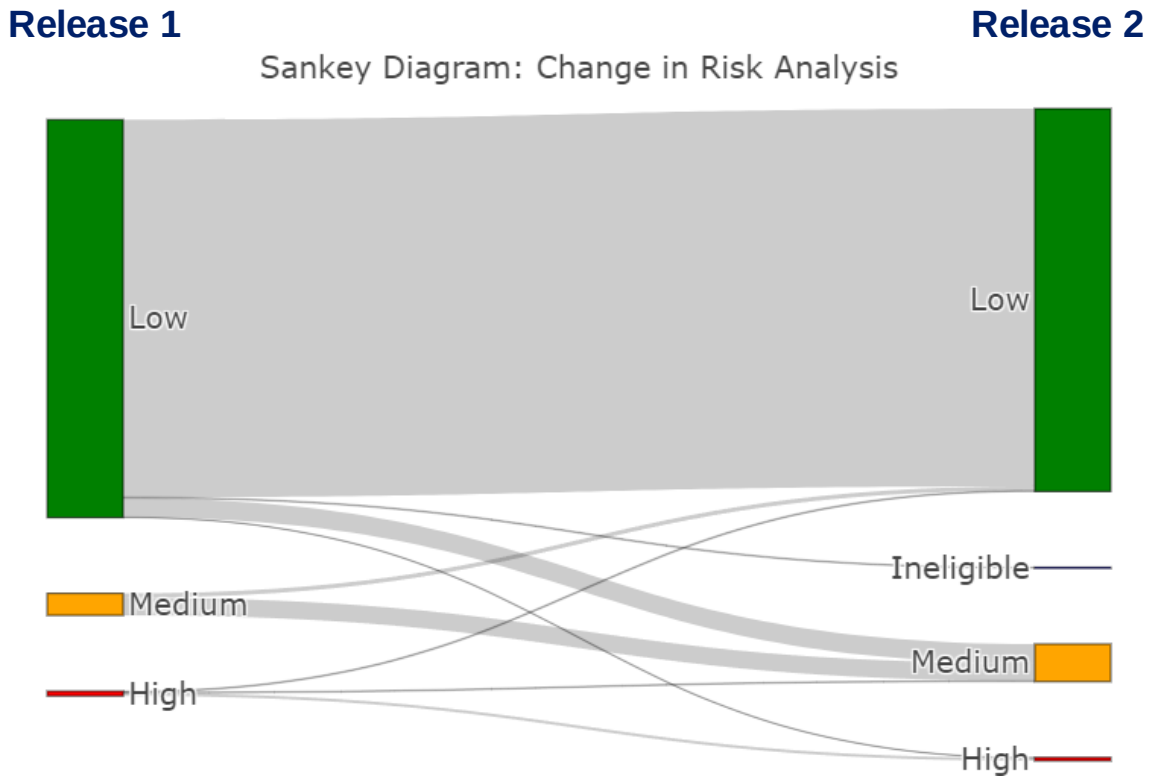
- Risk assessment unchanged: 836 **(93.7%)** packages



Risk Assessments Across Releases: Impact

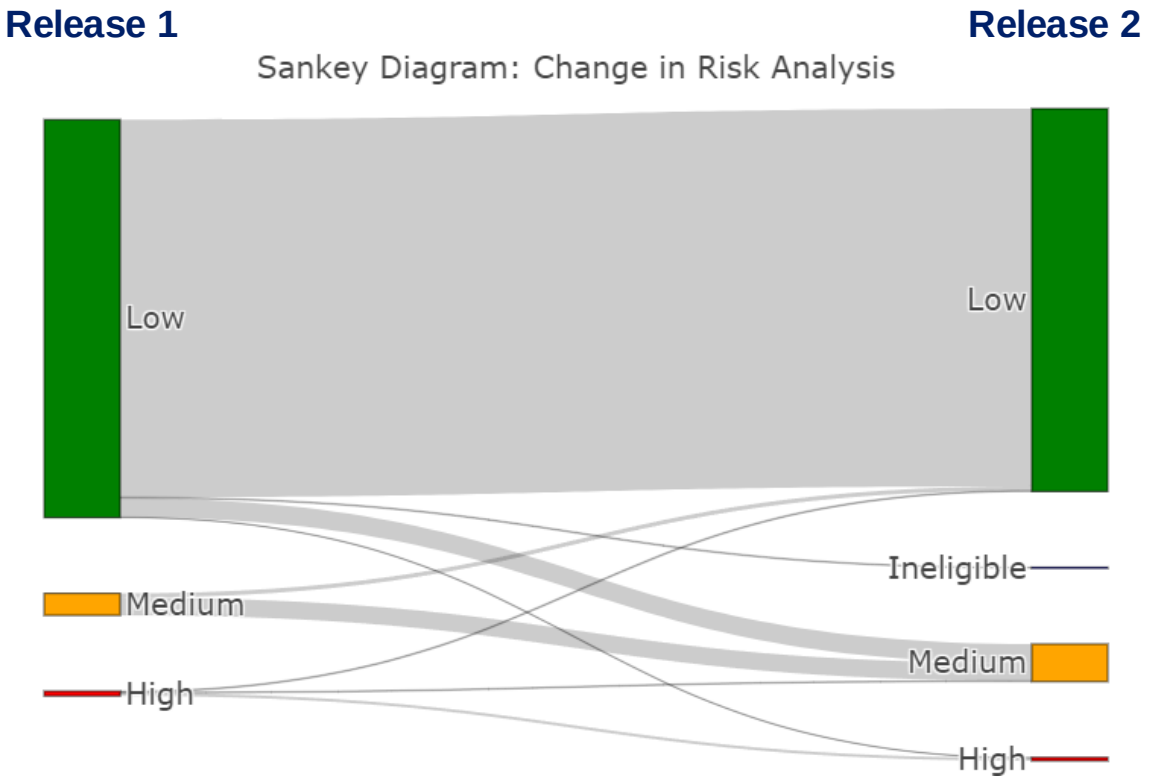
Total number of packages: 892

- Risk assessment unchanged: 836 **(93.7%)** packages
- Risk assessment changed: 56 **(6.3%)** packages



Risk Assessments Across Releases: Non-Impactful Changes

56 (6.3%) packages with changed risk assessments

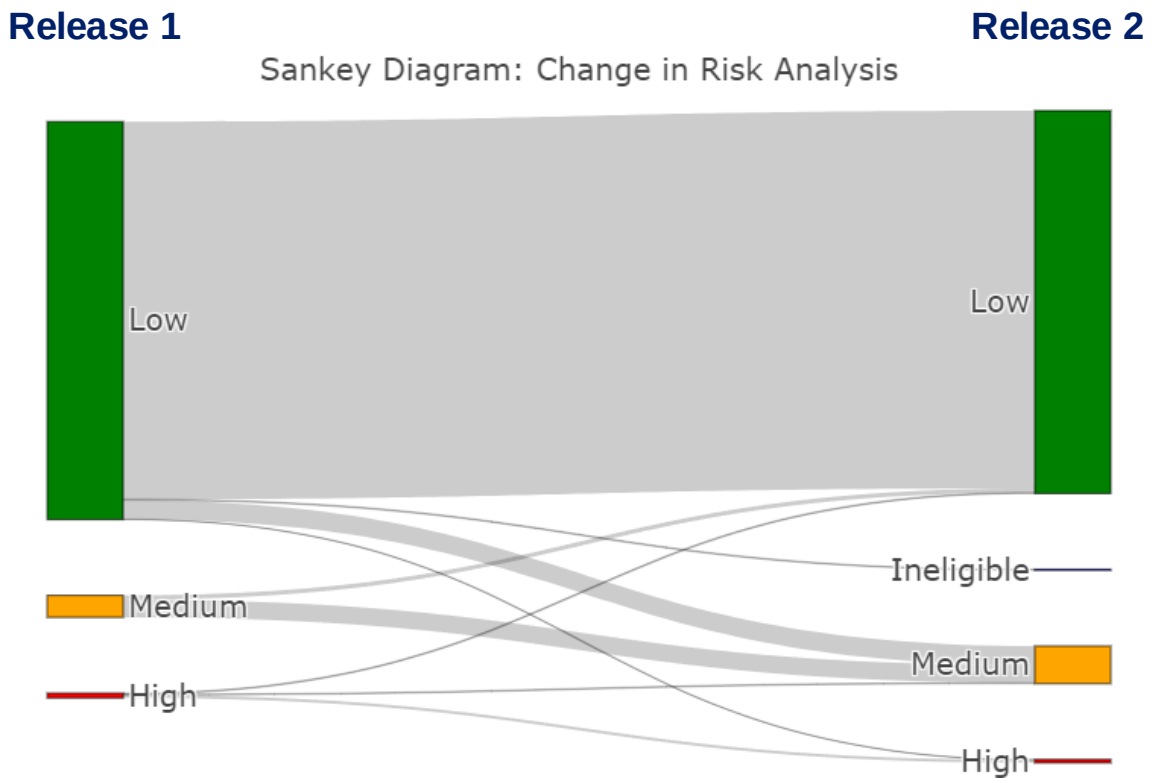


Risk Assessments Across Releases: Non-Impactful Changes

56 (6.3%) packages with changed risk assessments

Non-impactful changes

53 (94.6%) packages had non-impactful changes



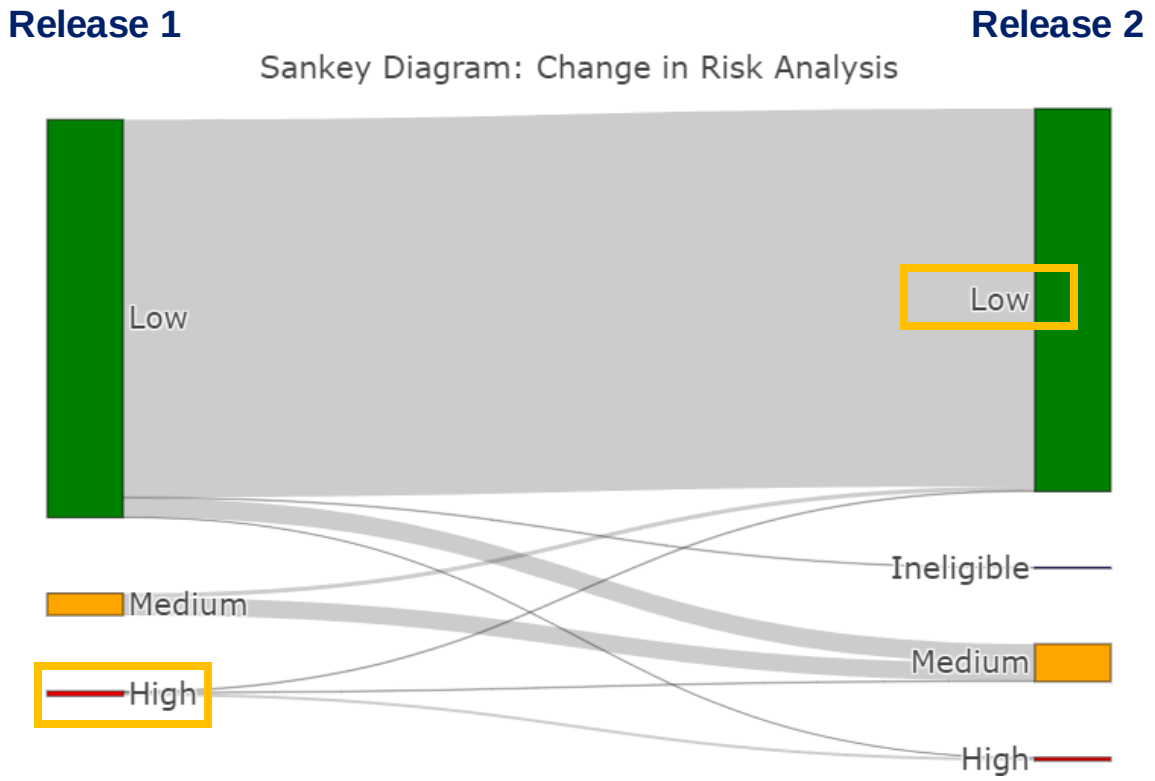
Risk Assessments Across Releases: Non-Impactful Changes

56 (6.3%) packages with changed risk assessments

Non-impactful changes

53 (94.6%) packages had non-impactful changes:

- High to Low Risk: **2 (3.6%)** packages



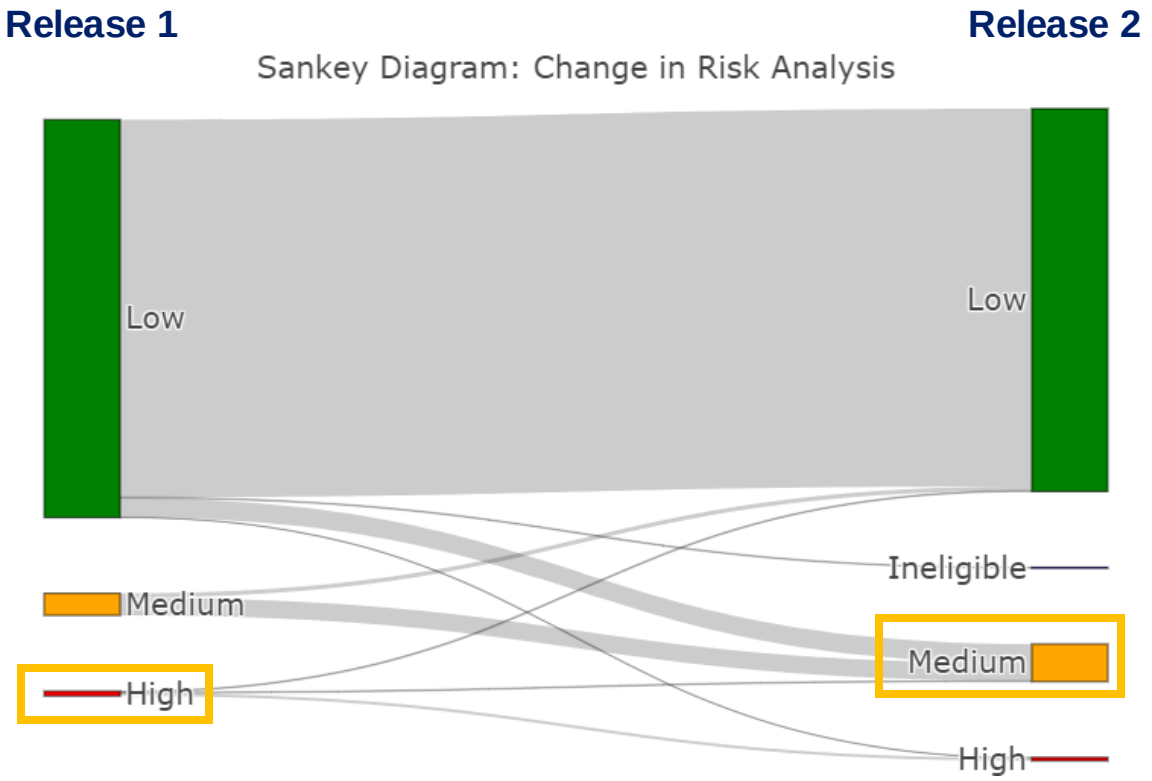
Risk Assessments Across Releases: Non-Impactful Changes

56 (6.3%) packages with changed risk assessments

Non-impactful changes

53 (94.6%) packages had non-impactful changes:

- High to Low Risk: 2 (3.6%) packages
- High to Medium Risk: **2 (3.6%)** packages



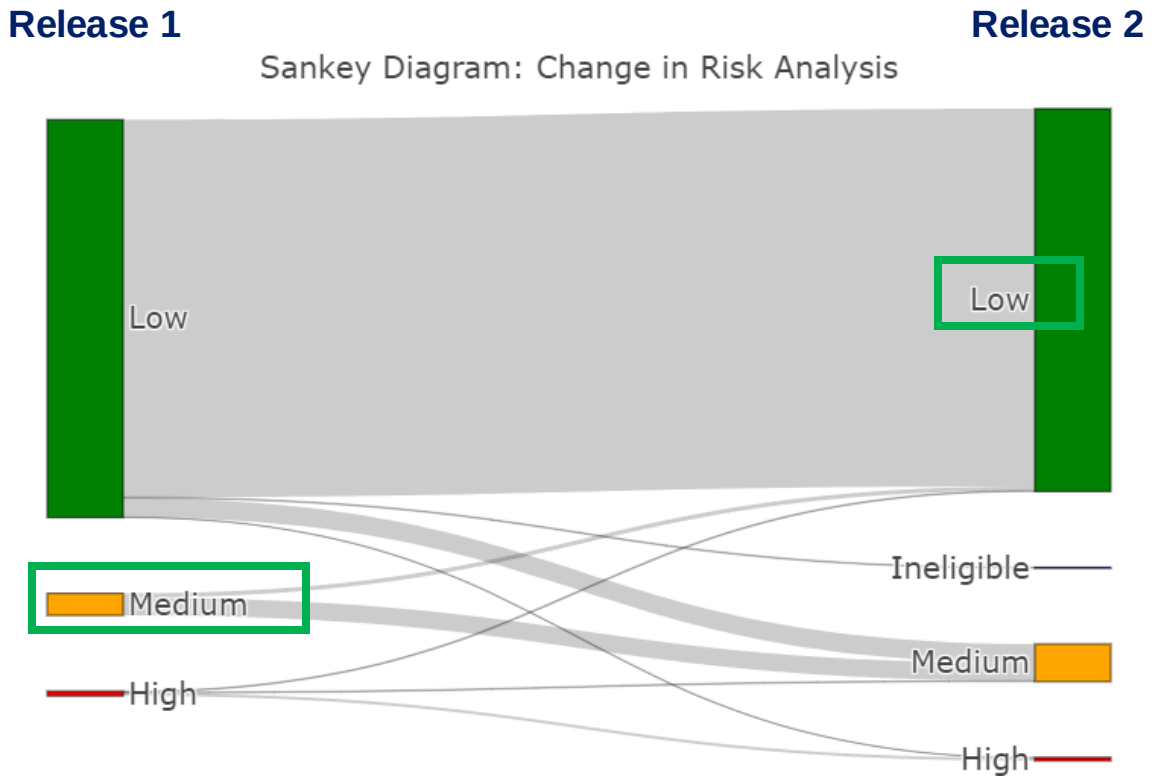
Risk Assessments Across Releases: Non-Impactful Changes

56 (6.3%) packages with changed risk assessments

Non-impactful changes

53 (94.6%) packages had non-impactful changes:

- High to Low Risk: 2 (3.6%) packages
- High to Medium Risk: 2 (3.6%) packages
- Medium to Low Risk: **9 (16.1%)** packages



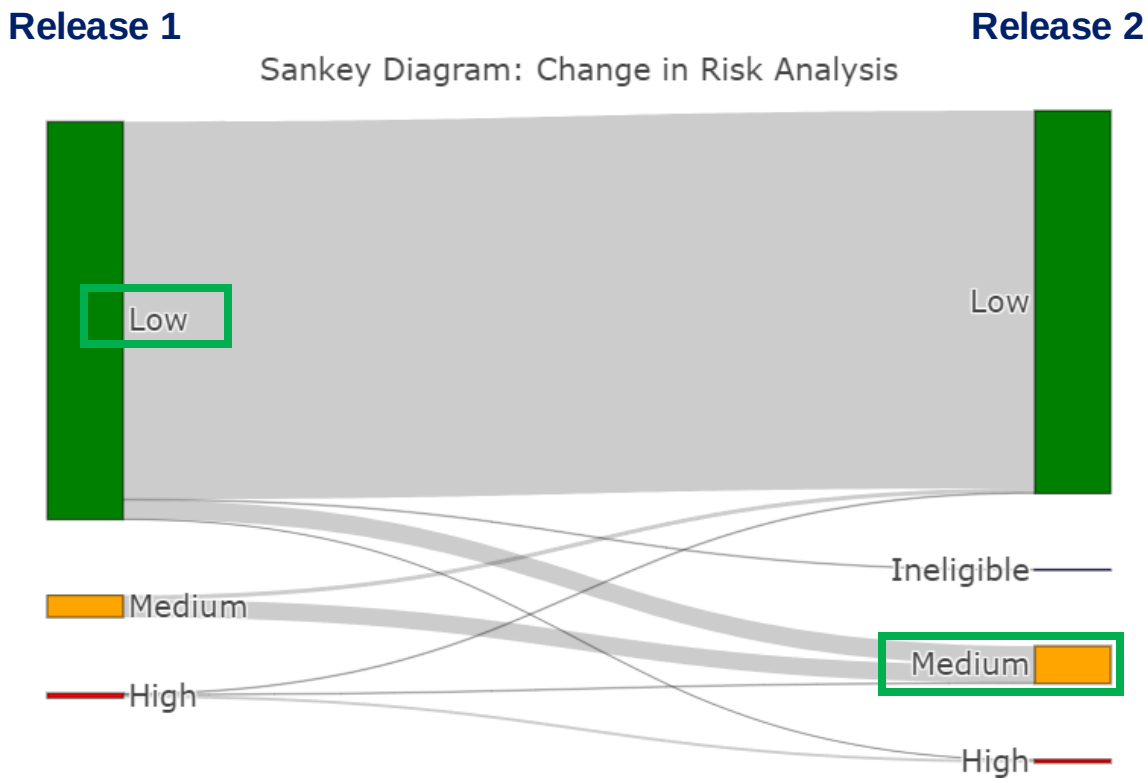
Risk Assessments Across Releases: Non-Impactful Changes

56 (6.3%) packages with changed risk assessments

Non-impactful changes

53 (94.6%) packages had non-impactful changes

- High to Low Risk: 2 (3.6%) packages
- High to Medium Risk: 2 (3.6%) packages
- Medium to Low Risk: 9 (16.1%) packages
- Low to Medium Risk: **40 (71.4%)** packages

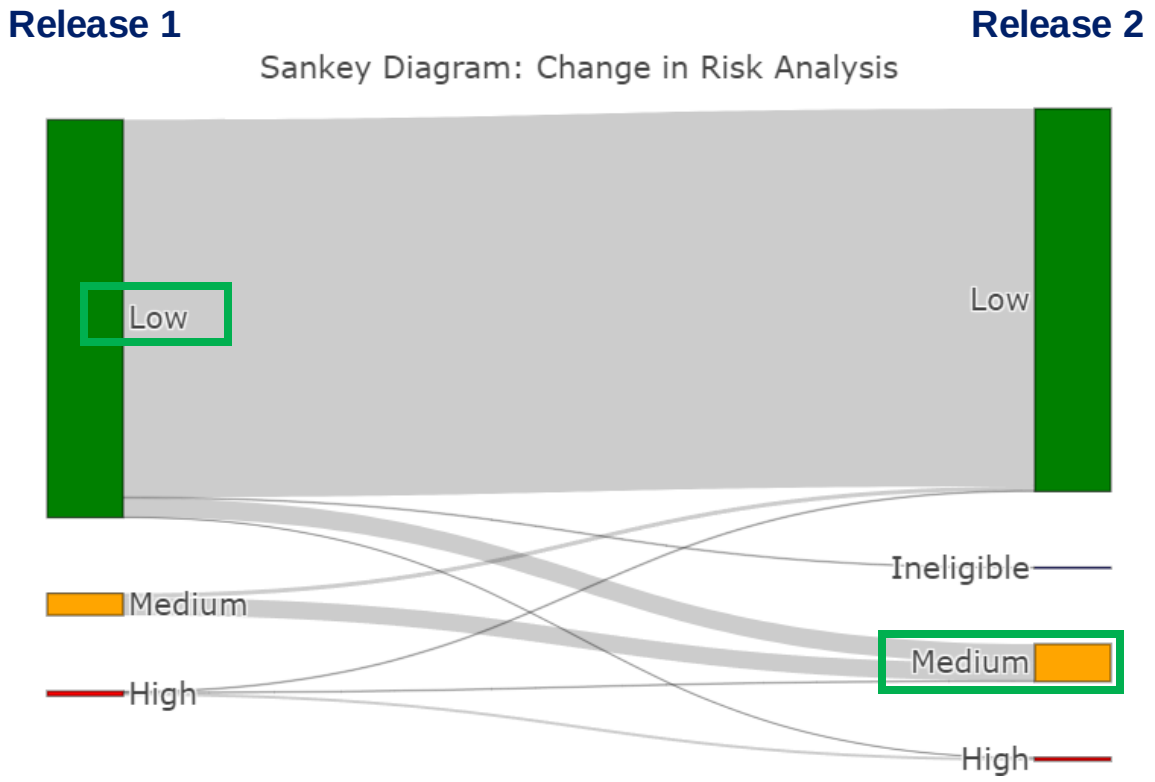


A Note on Download Rates

Low to Medium Risk: **40 (71.4%)** packages

Change in the risk assessment criteria

- Release 1 Low Risk Criteria: **1,000** downloads p/m
- Release 2 Low Risk Criteria: **50,000** downloads p/m

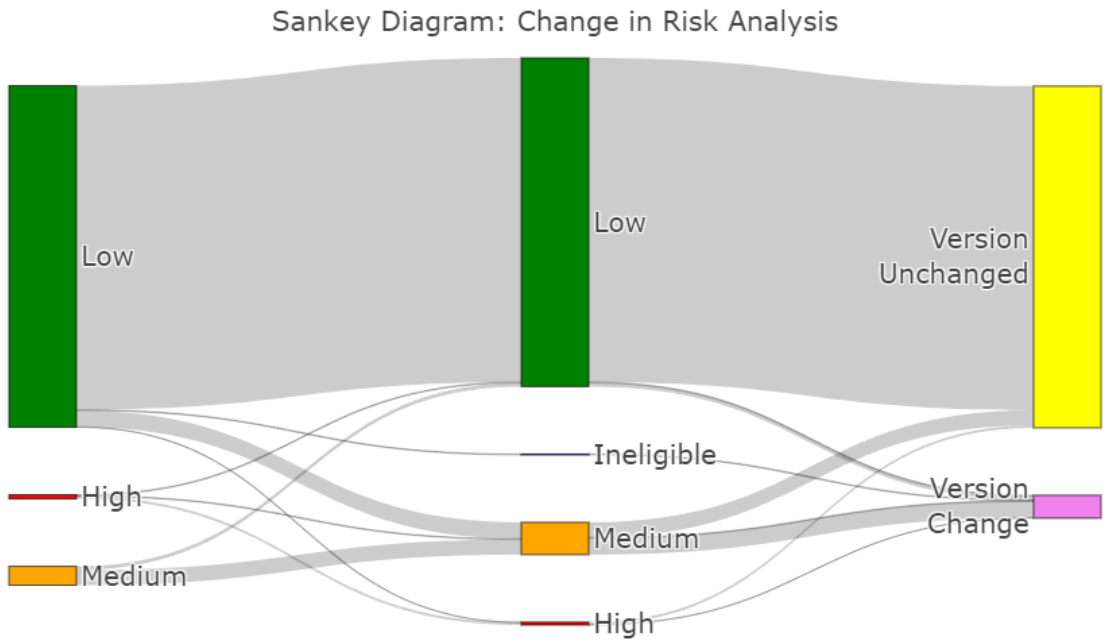


Risk Assessments Across Releases: Impactful Changes

56 (6.3%) packages with changed risk assessments

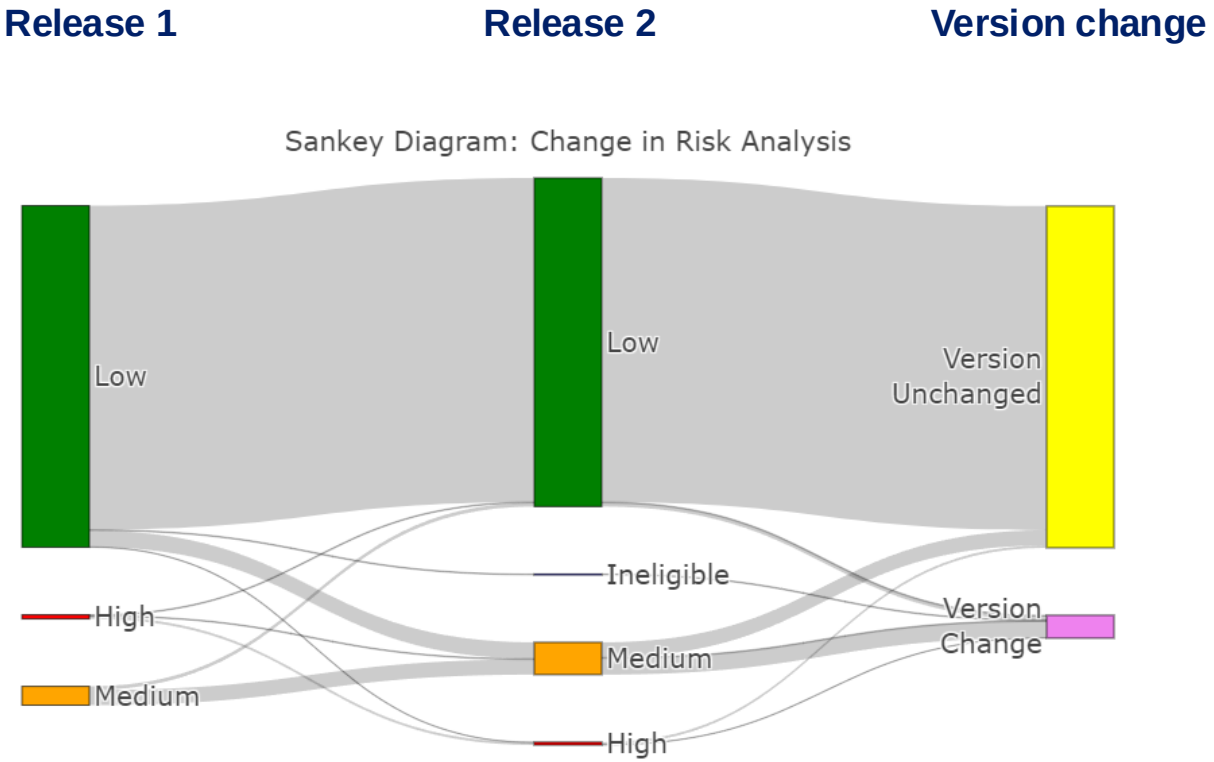
Release 1 **Release 2** **Version change**

Impactful changes



Risk Assessments Across Releases: Impactful Changes

56 (6.3%) packages with changed risk assessments



Impactful changes

3 (5.3%) packages had impactful changes

Risk Assessments Across Releases: Impactful Changes

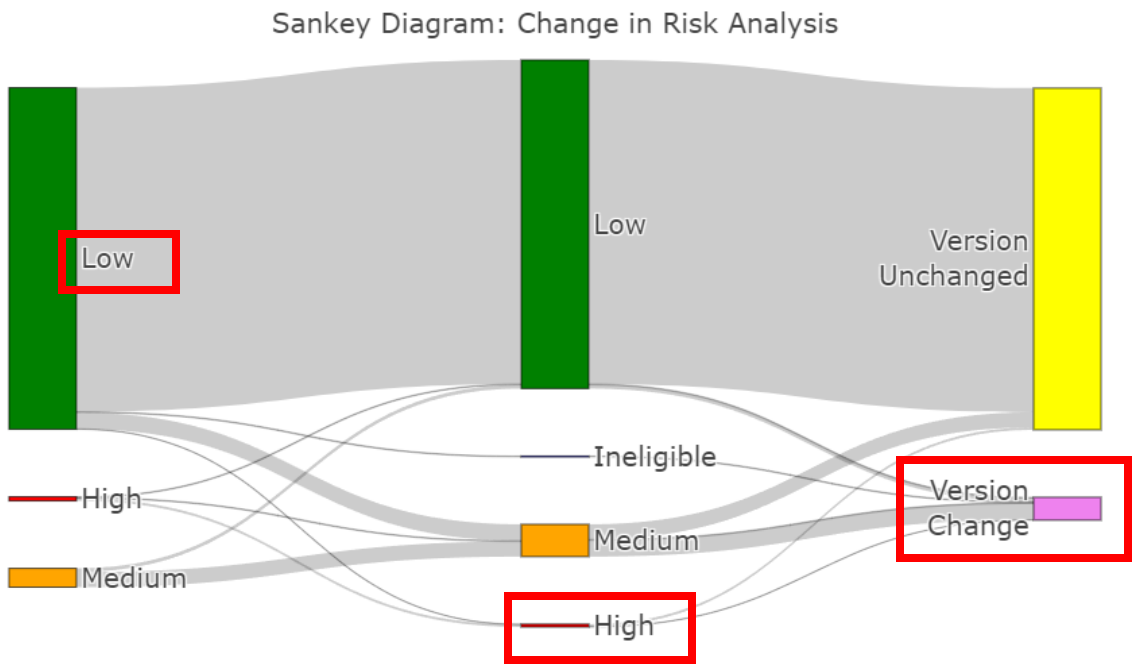
56 (6.3%) packages with changed risk assessments

Release 1 **Release 2** **Version change**

Impactful changes

3 (5.3%) packages had impactful changes

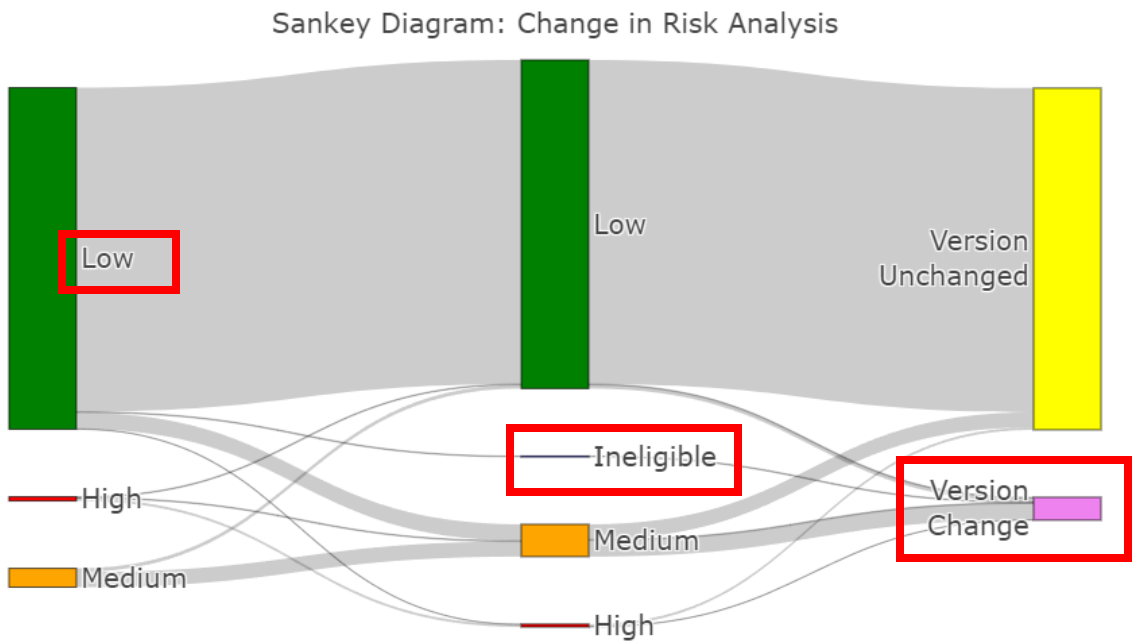
- Low to High Risk: **1 (1.8%)** package



Risk Assessments Across Releases: Impactful Changes

56 (6.3%) packages with changed risk assessments

Release 1 **Release 2** **Version change**



Impactful changes

3 (5.3%) packages had impactful changes

- Low to High Risk: 1 (1.8%) package
- Low Risk to Ineligible: **2 (3.6%)** packages

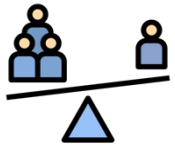
Risk Assessments Across Releases: Lessons Learned

Risk Assessments Across Releases: Lessons Learned

1. Are the risk assessments consistent across releases?

Risk Assessments Across Releases: Lessons Learned

1. Are the risk assessments consistent across releases?



Yes, for the vast majority

- 836 (93.7%) packages: no change to risk assessments
- 56 (6.3%) packages: change to risk assessments

Risk Assessments Across Releases: Lessons Learned

2. **Are the changes in the risk assessments impactful?**

Risk Assessments Across Releases: Lessons Learned

2. Are the changes in the risk assessments impactful?

56 (6.3%) packages had changed risk assessments

The majority are non-impactful

- **53** of these packages had non-impactful risk assessment changes
- Most changes in the risk assessment are changes within Low/Medium risk
- Small number of reductions in risk assessment (e.g. High → Low risk)

Risk Assessments Across Releases: Lessons Learned

2. Are the changes in the risk assessments impactful?

56 (6.3%) packages had changed risk assessments

The majority are non-impactful

- **53** of these packages had non-impactful risk assessment changes
- Most changes in the risk assessment are changes within Low/Medium risk
- Small number of reductions in risk assessment (e.g. High → Low risk)

Handful of impactful changes

- **3** of these packages had impactful risk assessment changes
- Risk assessment of 1 package increased from Low → High risk
- 2 packages deemed ineligible at Release 2

Risk Assessments Across Releases: Lessons Learned

3. What happens if the version of a package changes?

Risk Assessments Across Releases: Lessons Learned

3. What happens if the version of a package changes?

Only 3 packages (0.33% of total packages) have impactful changes

- 2 packages had version changes between releases
- Impact of package version being updated or not is unclear

Going Forward

Periodic Review

Periodically re-evaluate the risk assessment for installed packages

Whichever comes first:

Intervals of ~18 months

When the version of R is upgraded

Bríd Roberts
brid.roberts@novartis.com

Thank you

NOVARTIS | Reimagining Medicine