Detailed Guidance on Subgroup Analyses in Confirmatory Clinical Trials

1. Purpose of Subgroup Analyses

Subgroup analyses are conducted to explore whether the efficacy or safety of a treatment varies across different patient groups within a clinical trial. These groups may be defined by baseline characteristics such as age, sex, ethnicity, disease severity, genetic markers, or comorbidities. The goal is to identify potential heterogeneity in treatment effects that could inform clinical practice or regulatory decisions.

2. Planning and Pre-specification

- Subgroup analyses should be pre-specified in the study protocol and statistical analysis plan, including the rationale for selecting each subgroup.

- The number of subgroups and the criteria for their definition should be justified based on scientific, clinical, or regulatory relevance.

- Pre-specification helps control for multiplicity and reduces the risk of spurious findings due to post hoc data exploration.

3. Statistical Considerations

- Trials are typically powered to detect overall treatment effects, not effects within subgroups. Therefore, subgroup analyses often have limited statistical power.

- Interaction tests (e.g., treatment-by-subgroup interaction) should be used to formally assess whether treatment effects differ between subgroups.

- Multiple testing increases the risk of false positives; appropriate statistical adjustments (e.g., Bonferroni correction) should be considered if many subgroups are analyzed.

- Exploratory analyses of subgroups not pre-specified should be clearly identified as such and interpreted with caution.

4. Interpretation of Results

- Subgroup findings should be interpreted in the context of the overall trial results and the totality of evidence.

- Differences observed in subgroups may be due to chance, especially if the sample size is small or many subgroups are examined.

- Biological plausibility and consistency with external evidence (e.g., previous studies, mechanistic understanding) should be considered when evaluating subgroup effects.

- Subgroup analyses should not override the main trial conclusions unless there is strong, consistent, and credible evidence.

5. Reporting

- All planned and performed subgroup analyses should be transparently reported, including both positive and negative findings.

- Reports should include the rationale for subgroup selection, statistical methods used, and detailed results (effect estimates, confidence intervals, p-values).

- Forest plots and other graphical displays can help visualize subgroup effects and their consistency.

- The distinction between pre-specified and exploratory analyses should be clearly stated.

6. Regulatory Perspective

- Regulatory agencies require subgroup analyses to be pre-specified and scientifically justified.

- Subgroup findings may inform benefit-risk assessments, labeling, and recommendations for specific patient populations.

- Confirmatory evidence (e.g., replication in independent studies) is needed before making regulatory decisions based on subgroup effects.

- Agencies may request additional analyses or data to support subgroup claims.

7. Best Practices and Recommendations

- Limit the number of subgroups to those with strong scientific or clinical rationale.

- Use consistent definitions and methods across studies to facilitate comparison and meta-analysis.

- Engage with statisticians and regulatory experts early in trial design to ensure robust subgroup analysis plans.

- Interpret subgroup findings in the context of the broader evidence base, including mechanistic, epidemiological, and clinical data.

**General Guidance on Subgroup Analyses in Confirmatory Clinical Trials**

**Purpose:**  
Subgroup analyses are conducted to explore whether the treatment effect varies across different patient groups (e.g., age, sex, race, disease severity, biomarker status). These analyses help identify populations that may benefit more or less from the intervention.

**Key Principles:**

1. **Pre-specification:** Subgroups should be defined before the trial starts (pre-specified in the protocol or statistical analysis plan). Post hoc (exploratory) subgroup analyses are less reliable and should be interpreted with caution.
2. **Statistical Considerations:** Multiple subgroup analyses increase the risk of false positive findings (Type I error). Adjustments for multiplicity should be considered if formal claims are made. Interaction tests (testing for differences in treatment effect between subgroups) are preferred over separate analyses within each subgroup.
3. **Interpretation:** Subgroup findings should be considered hypothesis-generating unless strongly supported by biological rationale and statistical evidence. Consistency of treatment effect across subgroups strengthens the overall evidence. Apparent differences may be due to chance, especially in small subgroups.
4. **Reporting:** All planned and performed subgroup analyses should be transparently reported. Results should include estimates of treatment effect, confidence intervals, and interaction p-values.
5. **Regulatory Perspective:** Regulatory agencies generally require that the primary efficacy claim be based on the overall population. Subgroup claims may be considered if there is strong evidence and clinical relevance.

**Common Subgroups:**

* Demographics: Age, sex, race, ethnicity
* Disease characteristics: Severity, stage, histology
* Geographic region
* Biomarker status (e.g., PD-L1 expression)
* Baseline risk factors

**Best Practices:**

* Limit the number of subgroups to those with strong scientific or clinical justification.
* Use graphical displays (e.g., forest plots) to show consistency across subgroups.
* Avoid over-interpretation of isolated significant findings.

**Example Table: Subgroup Analysis Results**

| **Subgroup** | **N** | **Hazard Ratio (95% CI)** | **Interaction p-value** |
| --- | --- | --- | --- |
| Overall | 500 | 0.75 (0.60, 0.93) | - |
| Age <65 | 300 | 0.78 (0.61, 1.00) | 0.45 |
| Age ≥65 | 200 | 0.70 (0.50, 0.98) |  |
| Male | 250 | 0.80 (0.62, 1.03) | 0.60 |
| Female | 250 | 0.70 (0.52, 0.94) |  |
| PD-L1 Positive | 150 | 0.60 (0.40, 0.90) | 0.10 |
| PD-L1 Negative | 350 | 0.85 (0.65, 1.10) |  |

**Summary:**  
Subgroup analyses are valuable for understanding treatment effects in different patient populations, but must be carefully planned, analyzed, and interpreted. Regulatory agencies expect transparency, pre-specification, and appropriate statistical methods to avoid misleading conclusions.

Based on the extracted text from Amatya et al. (2021), here is a **summary guidance for subgroup analyses** in clinical trials, particularly in oncology, but applicable more broadly:

## Summary Guidance for Subgroup Analyses

### 1. ****Purpose and Importance****

* Subgroup analyses assess treatment effects in specific patient subsets within a trial, often defined by demographics, clinical, or molecular characteristics.
* They are important for interpreting trial results, regulatory decision-making, and informing product labeling.

### 2. ****Types of Subgroup Analyses****

* **Inferential:** Prespecified, adequately powered, and controlled for multiple testing; intended to establish efficacy in a subgroup.
* **Supportive:** Prespecified but not powered for formal testing; used to descriptively assess consistency of treatment effects.
* **Exploratory:** Not prespecified; used to generate hypotheses or gain further insight into treatment mechanisms.

### 3. ****Planning and Prespecification****

* Subgroups of interest should be prespecified in the study protocol, ideally based on biological rationale or prior evidence.
* Prespecification helps avoid data-driven, post hoc findings that may be spurious.

### 4. ****Statistical Considerations****

* Subgroup analyses often lack sufficient sample size (power) for reliable inference.
* Multiplicity (multiple comparisons) increases the risk of false-positive findings; appropriate statistical adjustments are necessary.
* Results from subgroup analyses should be interpreted with caution, especially if not prespecified or adequately powered.

### 5. ****Interpretation and Regulatory Use****

* Regulatory decisions consider the totality of evidence, including biological plausibility, clinical data, and external information.
* Subgroup findings may support narrowing or broadening an indication, but require strong justification (e.g., mechanistic rationale, consistent findings).
* Inconsistent or non-prespecified subgroup effects are generally considered hypothesis-generating rather than definitive.

### 6. ****Best Practices****

* Clearly define and justify subgroups before trial initiation.
* Ensure adequate sample size and statistical power if subgroup efficacy is a key objective.
* Adjust for multiple testing when conducting multiple subgroup analyses.
* Interpret subgroup findings in the context of overall trial results, biological plausibility, and clinical relevance.
* Report all planned and performed subgroup analyses transparently.

### 7. ****Cautions****

* Avoid over-interpreting subgroup findings, especially if based on small numbers or not prespecified.
* Recognize that subgroup analyses are more credible when supported by mechanistic understanding and consistent with other evidence.

**References for Further Reading:**

* ICH E9 Guideline on Statistical Principles for Clinical Trials
* FDA Guidance on Enrichment Strategies for Clinical Trials
* Amatya et al., Clin Cancer Res 2021;27:5753–6

**Summary of Guidance on Subgroup Analyses (Based on WR\_2019\_totext.txt):**

This document is a scoping review of guidance from key organisations on exploring, confirming, and interpreting subgroup effects of medical treatments. With the rise of personalized medicine, subgroup analyses are increasingly important but also challenging and prone to misuse, which can lead to false interpretations of treatment effects.

The review identified four types of key stakeholder organisations:

* Industry
* Health Technology Assessment (HTA) organisations
* Academic/non-profit research organisations
* Regulatory bodies

A total of 60 organisations were included, and 27 (45%) had relevant research guidance documents. The proportion of organisations providing guidance varied:

* 18% of industry organisations
* 64% of HTA organisations
* 38% of academic/non-profit organisations
* 57% of regulatory bodies

Most guidance documents (63%) mentioned challenges in subgroup analyses, such as false positives and ecological bias, with variations across organisation types. Statistical recommendations were less common (37%) and often limited to formal tests of interaction.

The review concludes that almost half of the organisations provide guidance on subgroup effect research, but there are large differences in the amount and detail of guidance. There is a need for more effort to translate research findings into practical guidelines and to harmonize guidance across organisations.

Subgroup analyses are complex due to the natural tension between detecting clinically relevant effects and avoiding incorrect claims. The document highlights the importance of integrating methodological research into practical decision-making and calls for reducing differences in guidance among organisations.

**Key Points:**

* Subgroup analyses are essential for personalized medicine but are methodologically challenging.
* Guidance varies widely between organisation types.
* Common challenges include false positives and ecological bias.
* Statistical guidance is often limited.
* Harmonization and translation of research into practice are needed.

Based on the summary and the content of the document, here is **practical guidance for conducting subgroup analyses** in medical research:

## Practical Guidance for Subgroup Analyses

### 1. ****Define Subgroups A Priori****

* Identify and justify subgroups before data analysis, based on biological plausibility, clinical relevance, or previous evidence.
* Avoid data-driven or post hoc subgroup definitions, as these increase the risk of spurious findings.

### 2. ****Limit the Number of Subgroup Analyses****

* Restrict the number of subgroup analyses to those that are clinically meaningful and pre-specified.
* Multiple testing increases the risk of false positive results.

### 3. ****Use Appropriate Statistical Methods****

* Employ formal statistical tests for interaction to assess whether treatment effects differ between subgroups.
* Avoid interpreting subgroup effects based solely on within-group comparisons.

### 4. ****Interpret Results with Caution****

* Recognize that subgroup analyses are often underpowered and may yield unreliable results.
* Consider the totality of evidence, including consistency with other studies and biological plausibility.

### 5. ****Report Subgroup Analyses Transparently****

* Clearly state which subgroup analyses were pre-specified and which were exploratory.
* Report all conducted subgroup analyses, regardless of statistical significance, to avoid selective reporting.

### 6. ****Beware of Common Pitfalls****

* Be cautious of ecological bias and over-interpretation of findings.
* Understand that statistically significant interactions may occur by chance, especially with many subgroups.

### 7. ****Consult Guidance and Standards****

* Follow relevant guidelines from regulatory bodies, HTA agencies, and academic organisations.
* Stay updated with methodological research on subgroup analysis.

### 8. ****Integrate Findings into Decision-Making Carefully****

* Use subgroup findings to inform, but not dictate, clinical or policy decisions unless robust and replicated evidence exists.

**References:**

* BMJ Open 2019;9:e028751. doi:10.1136/bmjopen-2018-028751
* CONSORT Statement (for reporting subgroup analyses in clinical trials)
* Regulatory and HTA guidance documents

Here is **practical guidance for subgroup analyses in clinical trials**, based on best practices and the themes from your document:

## Subgroup Analysis Guidance for Clinical Trials

### 1. ****Pre-specify Subgroups in the Protocol****

* Define subgroups before the trial begins, based on clinical relevance, biological rationale, or prior evidence.
* Document subgroup hypotheses and analysis plans in the trial protocol and statistical analysis plan.

### 2. ****Limit the Number of Subgroups****

* Restrict analyses to a small number of clinically justified subgroups to minimize the risk of false positives.
* Avoid exploratory, data-driven subgroup analyses unless clearly labeled as such.

### 3. ****Use Formal Interaction Tests****

* Assess differences in treatment effect between subgroups using statistical tests for interaction (e.g., including an interaction term in regression models).
* Do not rely solely on separate analyses within each subgroup.

### 4. ****Interpret with Caution****

* Recognize that subgroup analyses are often underpowered and may yield unreliable or non-reproducible results.
* Consider the plausibility and consistency of findings with external evidence.

### 5. ****Transparent Reporting****

* Clearly distinguish between pre-specified and post hoc (exploratory) subgroup analyses in publications.
* Report all subgroup analyses conducted, regardless of statistical significance, to avoid selective reporting.

### 6. ****Adjust for Multiple Comparisons****

* Consider the increased risk of type I error due to multiple subgroup analyses.
* Use appropriate statistical adjustments or interpret findings as hypothesis-generating.

### 7. ****Avoid Common Pitfalls****

* Do not over-interpret statistically significant findings from small subgroups.
* Beware of ecological bias and confounding.

### 8. ****Follow CONSORT Guidelines****

* Adhere to the CONSORT extension for reporting subgroup analyses in randomized trials.
* Provide rationale, methods, and results for each subgroup analysis.

### 9. ****Consult Regulatory and HTA Guidance****

* Review relevant guidance from regulatory agencies (e.g., FDA, EMA) and HTA bodies for specific requirements and recommendations.

**Example Checklist for Subgroup Analyses in Clinical Trials:**

1. Are subgroups pre-specified in the protocol?
2. Is the rationale for each subgroup clear and evidence-based?
3. Are formal interaction tests used?
4. Are results interpreted in the context of overall trial findings and external evidence?
5. Are all subgroup analyses transparently reported?
6. Is the risk of false positives due to multiple testing addressed?
7. Are findings used to inform, not dictate, clinical decisions unless robustly supported?

**References:**

* BMJ Open 2019;9:e028751. doi:10.1136/bmjopen-2018-028751
* CONSORT Statement: [CONSORT Subgroup Analysis Extension](http://www.consort-statement.org/extensions?content=subgroup-analysis)
* FDA, EMA, and HTA guidance documents

Here is a structured summary of subgroup analysis results for KEYTRUDA (pembrolizumab) based on the extracted text from the prescribing information:

### 1. ****Non-Small Cell Lung Cancer (NSCLC)****

* **PD-L1 Expression Subgroups (TPS ≥50%, ≥20%, ≥1%)**:
  + Efficacy outcomes (Overall Survival [OS], Progression-Free Survival [PFS], Objective Response Rate [ORR]) were analyzed in subgroups based on PD-L1 Tumor Proportion Score (TPS).
  + The main efficacy outcome was OS in the TPS ≥50% subgroup, TPS ≥20% subgroup, and the overall population with TPS ≥1%.
  + Results in the TPS ≥20% subgroup were intermediate between those with TPS ≥1% and TPS ≥50%.
  + In a pre-specified exploratory subgroup analysis for patients with TPS 1-49%, the median OS was 13 months.
* **First-line and Previously Treated Populations**:
  + Subgroup analyses were performed for patients receiving KEYTRUDA as a single agent or in combination with chemotherapy, stratified by PD-L1 status and prior therapy.

### 2. ****Head and Neck Squamous Cell Carcinoma (HNSCC)****

* **PD-L1 Combined Positive Score (CPS) Subgroups (CPS ≥20, CPS ≥1)**:
  + Statistically significant improvement in OS for subgroups with CPS ≥1 and CPS ≥20.
  + KEYNOTE-048 trial showed improved OS for KEYTRUDA in patients with CPS ≥1 compared to cetuximab plus chemotherapy.
  + Exploratory subgroup analysis for CPS 1-19 showed a median OS of 10 months.

### 3. ****Other Tumor Types****

* **Microsatellite Instability-High (MSI-H) or Mismatch Repair Deficient (dMMR) Tumors**:
  + Subgroup analyses included adult and pediatric patients with unresectable or metastatic MSI-H/dMMR solid tumors, including colorectal cancer (CRC).
  + Efficacy and safety were evaluated in these biomarker-defined subgroups.
* **HER2-Positive Gastric or GEJ Adenocarcinoma**:
  + Subgroup analysis for patients with HER2-positive tumors receiving KEYTRUDA in combination with trastuzumab and chemotherapy.
* **Cervical Cancer**:
  + Subgroup analysis for patients with tumors expressing PD-L1 (CPS ≥1).
* **Triple-Negative Breast Cancer (TNBC)**:
  + Subgroup analysis for patients with locally recurrent unresectable or metastatic TNBC whose tumors express PD-L1 (CPS ≥10).

### 4. ****Safety Analyses****

* **Dose and Schedule Subgroups**:
  + Safety profiles were similar across different dosing regimens (e.g., 2 mg/kg vs. 10 mg/kg, every 2 weeks vs. every 3 weeks), so pooled analyses were presented.
* **Exploratory Subgroups**:
  + Analyses by age, sex, race, prior therapy, and other baseline characteristics were performed, but no new safety signals were identified in these subgroups.

### 5. ****Other Notable Subgroup Analyses****

* **Adjuvant and Neoadjuvant Settings**:
  + Subgroup analyses for disease-free survival (DFS) in patients who did or did not receive adjuvant chemotherapy.
* **Exploratory and Pre-specified Analyses**:
  + Multiple trials reported exploratory and pre-specified subgroup analyses for efficacy and safety endpoints, often stratified by biomarker status, prior treatment, and demographic factors.

**Summary:**  
KEYTRUDA’s efficacy and safety have been evaluated in multiple subgroups defined by tumor type, biomarker status (PD-L1, CPS, MSI-H/dMMR), prior therapy, and demographic characteristics. The most consistent findings are that higher PD-L1 expression is associated with greater benefit, and that efficacy is maintained across a range of tumor types and patient subgroups. Safety profiles are generally consistent across subgroups and dosing regimens.

Here is a focused summary of subgroup analysis results for KEYTRUDA (pembrolizumab) in Non-Small Cell Lung Cancer (NSCLC) based on the prescribing information:

### NSCLC Subgroup Analysis Summary

#### 1. ****PD-L1 Expression Subgroups****

* **KEYTRUDA efficacy was analyzed in subgroups based on PD-L1 Tumor Proportion Score (TPS):**
  + **TPS ≥50%:**
    - This subgroup showed the greatest benefit from KEYTRUDA, with improved overall survival (OS), progression-free survival (PFS), and objective response rate (ORR) compared to chemotherapy.
  + **TPS ≥20%:**
    - Efficacy results in this subgroup were intermediate between those with TPS ≥1% and TPS ≥50%.
  + **TPS ≥1% (overall population):**
    - KEYTRUDA provided a survival benefit compared to chemotherapy in the overall population with TPS ≥1%.
  + **TPS 1-49%:**
    - In a pre-specified exploratory subgroup analysis, the median OS was 13 months.

#### 2. ****First-Line and Previously Treated Populations****

* **First-Line Treatment:**
  + KEYTRUDA, alone or in combination with chemotherapy, was evaluated in patients with metastatic NSCLC, stratified by PD-L1 status.
  + The main efficacy outcome was OS in the TPS ≥50% subgroup, TPS ≥20% subgroup, and the overall population with TPS ≥1%.
  + Additional outcomes (PFS, ORR) were also reported by these subgroups.
* **Previously Treated Patients:**
  + KEYTRUDA as a single agent was evaluated in patients with metastatic NSCLC whose tumors express PD-L1 (TPS ≥1%) and who had disease progression on or after platinum-containing chemotherapy.
  + Patients with EGFR or ALK genomic tumor aberrations were required to have disease progression on FDA-approved therapy for these aberrations before receiving KEYTRUDA.

#### 3. ****Other Subgroup Analyses****

* **Safety:**
  + Safety profiles were similar across different dosing regimens (2 mg/kg vs. 10 mg/kg, every 2 weeks vs. every 3 weeks), so pooled safety analyses were presented.
* **Exploratory Subgroups:**
  + Analyses by age, sex, race, prior therapy, and other baseline characteristics did not reveal new safety signals or significant differences in efficacy.

#### 4. ****Summary Table and Figures****

* The prescribing information includes tables and Kaplan-Meier curves summarizing efficacy results in the TPS ≥50% subgroup and the overall population with TPS ≥1%.
* The results consistently show that higher PD-L1 expression (TPS ≥50%) is associated with greater benefit from KEYTRUDA.

**Key Point:**  
The benefit of KEYTRUDA in NSCLC is most pronounced in patients with higher PD-L1 expression (TPS ≥50%), but efficacy is also observed in patients with TPS ≥1%. Subgroup analyses support the use of KEYTRUDA across a range of patient characteristics, with consistent safety profiles.

Draft summary of above text –

## Guidance on Subgroup Analyses in Confirmatory Clinical Trials

### 1. Purpose of Subgroup Analyses

* Subgroup analyses explore whether treatment efficacy or safety varies across different patient groups (e.g., age, sex, ethnicity, disease severity, genetic markers, comorbidities).
* The goal is to identify potential heterogeneity in treatment effects to inform clinical practice or regulatory decisions.

### 2. Planning and Pre-specification

* Subgroup analyses should be pre-specified in the study protocol and statistical analysis plan, with clear rationale for each subgroup.
* The number and definition of subgroups should be justified based on scientific, clinical, or regulatory relevance.
* Pre-specification helps control for multiplicity and reduces the risk of spurious findings from post hoc exploration.

### 3. Statistical Considerations

* Clinical trials are usually powered for overall effects, not for subgroups; thus, subgroup analyses often lack statistical power.
* Use interaction tests (e.g., treatment-by-subgroup interaction) to formally assess differences between subgroups.
* Multiple testing increases false positive risk; apply appropriate statistical adjustments (e.g., Bonferroni correction) if many subgroups are analyzed.
* Clearly identify and interpret exploratory (not pre-specified) subgroup analyses with caution.

### 4. Interpretation of Results

* Interpret subgroup findings in the context of overall trial results and the totality of evidence.
* Differences in subgroups may be due to chance, especially with small sample sizes or many subgroups.
* Consider biological plausibility and consistency with external evidence when evaluating subgroup effects.
* Subgroup analyses should not override main trial conclusions unless evidence is strong, consistent, and credible.

### 5. Reporting

* Transparently report all planned and performed subgroup analyses, including both positive and negative findings.
* Reports should include rationale for subgroup selection, statistical methods, and detailed results (effect estimates, confidence intervals, p-values).