Statistical Analysis Plan

Administrative information	
Study title	Thoracentesis to alleviate pleural effusion in acute heart failure
Study acronym	The TAP-IT trial
Clinicaltrials.gov ID:	NCT05017753.
SAP version	SAP version: 1.1 Date: 09-Nov-2023
SAP Revision history	Non
Protocol version	Study protocol version 2.0, January 7th 2022 (was approved as an amendment to the original study protocol version 1.0 dated December 21st 2021)
Approvals	The Capital Region of Denmark Scientific Ethical Committee (H-20060817) Knowledge Centre for Data Reviews (P-2021-149). The study does not require registration and monitoring from the Danish Medicines Agency (case number 2020031478).
Trial Principal Investigators	Jens Jakob Thune, MD, PhD, associate professor Department of Cardiology, Copenhagen University Hospital - Bispebjerg and Frederiksberg, Copenhagen, Denmark And Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark
Authors	Signe Glargaard, MD, PhD fellow Jakob Hartvig Thomsen, MD, PhD Department of Cardiology, Copenhagen University Hospital - Bispebjerg and Frederiksberg, Copenhagen, Denmark

Role and responsibilities	Signe Glargaard drafted the SAP. Jakob Hartvig Thomsen consulted on the SAP Jens Jakob Thune consulted and approved the final SAP.
Senior Statistical Responsible	Jens Jakob Thune
SAP deadline	The SAP will be finalised before the inclusion of the last patient.

Introduction

Background and rationale

Pleural effusion is a common finding in patients admitted with acute heart failure and congestion, and patients with larger heart failure-related pleural effusion necessitating therapeutic thoracentesis (from here on only referred to as thoracentesis) have high mortality rates [1,2]. The current treatment options for heart failure-related pleural effusions are decongestion with diuretics in combination with initiation or adjustment of guideline-directed medical therapy (GDMT) for the underlying heart failure condition, and for larger effusions, thoracentesis is an option. However, thoracentesis in heart failure-related pleural effusions is controversial. There is no evidence from randomised trials, and there is no recommendation in international guidelines [3,4]. The timeliness and benefits of thoracentesis have not previously been studied.

Study objectives

Primary objective

To investigate if a strategy of referral to up-front thoracentesis in addition to pharmacological therapy compared with pharmacological therapy alone increases the number of days alive outside of the hospital during the following 90 days in patients with pleural effusion due to acute heart failure and left ventricular ejection fraction (LVEF) ≤45%.

Research hypothesis

A strategy of referring patients with heart failure-related pleural effusion to upfront thoracentesis by pigtail catheter insertion increases the number of days alive outside of the hospital during the following 90 days.

Secondary objectives

To determine the effectiveness of a strategy of referral to upfront thoracentesis in addition to pharmacological therapy compared with pharmacological therapy alone on admission duration, dose of

diuretics, weight loss, mortality, complications, and patient-reported outcomes such as patient satisfaction and quality of life.

Study Methods

Study design

The TAP-IT trial is an investigator-initiated, pragmatic, multicentre, randomised, parallel-group, open-label, controlled trial. Treatment allocation is 1:1. Participants in the intervention group are assigned to referral to upfront pigtail catheter insertion (thoracentesis) combined with standard pharmacological therapy (diuretics and GDMT). Participants in the control group are assigned to standard pharmacological therapy alone (comparator). Physicians, investigators, and participants are not blinded to the result of the randomisation.

Randomisation

Participants are randomised with an internet-based electronic case report form and randomisation program. Randomisation is stratified according to site and whether participants are treated with oral anticoagulation therapy (regardless of type: Direct-acting oral anticoagulant or Vitamin K antagonist vs. no anticoagulation) with alternating block sizes to reduce predictability.

Sample Size

For estimation of sample size, we assumed a t-test of superiority and found a total of 126 participants required to detect a difference of three days in the primary endpoint with an α of 0.05 and a power of 90%. This assumes, that participants assigned to a strategy with referral to up-front thoracentesis in addition to pharmacological therapy will have 85 days alive and not hospitalised during the 90 days after randomisation, while participants assigned to pharmacological therapy alone will have 82 days, with a shared standard deviation of 5 days, and in-hospital mortality of 5% in both groups [5–7]. For analysis of the primary endpoint, we will assess the Mann-Whitney parameter for days alive without hospitalisation during the 90 days following randomisation, using the Wilcoxon-Mann-Whitney-test [30]. The loss of power with the Wilcoxon rank sum test compared to the t-test is often limited if distributions are normal, and when normality is violated, the Wilcoxon rank sum test can be three or four times more powerful than the independent samples t-test [31,32]. The power of 90% will allow for a possible small loss of power from using the Wilcoxon-Mann-Whitney test.

Statistical interim analysis and stopping guidance

No statistical interim analysis will be performed. There is no stopping guidance since the intervention and comparator are already standard treatments with well-known complications and adverse effects.

Adherence and protocol deviations and Randomization errors

The intervention and comparator in this pragmatic trial are two treatment strategies: referral to up-front thoracentesis in addition to pharmacological therapy compared with pharmacological therapy alone. Compliance with the protocol is defined as the participant in the intervention arm being referred to upfront thoracentesis by the treating physician (scheduling when the thoracentesis will be performed). There is the possibility that the physician performing the procedure does not find enough pleural fluid to safely perform the thoracentesis. This case will not be considered a violation of the protocol. In the control group, participants receiving pharmacological treatment alone for five days after randomisation are considered compliant with the protocol.

Major protocol deviation is defined as patients randomised despite violating eligibility criteria. This is considered a randomisation error. Such cases will remain in the trial and have data collected. The ineligible participant will be included in the full analysis dataset because it is our perception that exclusions cannot be made in an objective and unbiased manner due to the unblinded nature of the trial, with a risk of introducing bias post-randomization [8]. The allocated treatment strategy is stopped if the reason for ineligibility means that commencing or continuing treatment is inappropriate or potentially unsafe (e.g. participants randomised despite an indication for diagnostic thoracentesis must be referred to thoracentesis; participants randomised despite a spontaneous bleeding disorder and therefore cannot undergo thoracentesis). Such cases will be analysed according to the allocated treatment arm in accordance with the intention-to-treat principle.

Trial population

Eligibility criteria are stated in the study protocol version 2.0 (January 7th 2022). The criteria have been updated with minor changes from version 1.0 (December 21st 2021). Primarily one additional exclusion criterion was added “planned or expected admission >10 days for other condition than heart failure”, and two exclusion criteria were clarified: “Known or suspected malignant disease” was changed to “clinically indicated diagnostic thoracentesis (e.g. suspected malignant aetiology)” and “eGFR<15ml/min/1.73m^2 or acute renal failure” was changed to “eGFR<15ml/min/1.73m^2 or dialysis treatment”. Eleven participants were included under the criteria in protocol version 1.0. The impact of the changes in eligibility criteria is

believed to be minor, and we do not see the need for stratification of the population based on whether they were included before or after the implementation of protocol version 2.0.

The following will be reported.

- The number of patients screened for eligibility with an explanation for the exclusion
- The number of patients randomised
- The number of participants allocated to each group.
- The number of participants that followed the treatment strategy according to the protocol (patients in the intervention arm being referred to upfront thoracentesis by the treating physician (scheduling when the thoracentesis will be performed) and patients in the control arm who are not referred to thoracentesis within five days after randomisation)
- The number of participants that received the planned treatment according to the allocated treatment strategy (thoracentesis performed on participants in the interventional group and thoracentesis not performed on participants in the control group).
- The number of participants in the intervention group that did not have thoracentesis performed and the number of participants in the control group that had thoracentesis.
- The number of participants that withdrew from the trial or were lost to follow-up with an explanation. The level of consent withdrawal will be reported ("consent to continued data collection and questionnaire follow-up", "consent to continued data collection but no questionnaire follow-up", "complete – no questionnaire follow-up or data collection")
- The number of participants analysed for the primary outcome according to the intention-to-treat principle.

An overview of the trial population will be provided in a CONSORT flow diagram (presented as Figure 1 in the final publication)

Statistical methods

Analysis datasets

The primary analysis will be on the **full analysis dataset,** including all participants from whom written informed consent is obtained and randomised. Subjects will be analysed according to their allocated treatment group according to the intention-to-treat principle.

Exploratory analyses will be performed on the per-protocol and the as-treated datasets as defined below. The **per-protocol dataset** includes participants that followed the treatment strategy according to the

protocol (patients in the intervention arm being referred to upfront thoracentesis by the treating physician (scheduling when the thoracentesis will be performed) and patients in the control arm who are not referred to thoracentesis within five days after randomisation). The **as-treated dataset** is defined as participants who received the planned treatment according to the allocated treatment strategy (thoracentesis performed on participants in the interventional group and thoracentesis not performed on participants in the control group).

Statistical principles

All outcomes are tested for superiority. A two-sided p-value <0.05 is considered statistically significant, and a 95% confidence interval (CI) is reported for all effect estimates. No multiplicity adjustment will be applied.

Timing of final analysis

All outcomes are analysed collectively after the 90-day follow-up of the last patient when data collection is completed.

Baseline characteristics

Baseline patient characteristics and demographics will be reported for each group (Table 1 in the final publication).

Table 1: Baseline characteristics	
Variable	Data type
Demographics	
Age (years)	Continuous
Sex (female) %(n)	Categorical
Race (white)	Categorical
Included at specialised referral centre %(n)	Categorical
Symptoms and physical examination	
Height (cm)	Continuous
Weight (kg)	Continuous
BMI (kg/m2)	Continuous
Rales %(n)	Categorical
Peripheral oedema %(n)	Categorical
Orthopnoea %(n)	Categorical
Ascites %(n)	Categorical

NYHA class %(n) II III IV	Categorical, ordinal
Fatigue %(n)	Categorical
The CSHA clinical frailty scale score [9,10] Score<4 = frail %(n) Score>4 = not frail %(n)	Categorical, ordinal
Heart failure characteristics	
LVEF (%)	Continuous
Time since HF diagnosis (months)	Continuous
New-onset HF %(n)	Categorical
Ischemic aetiology %(n)	Categorical
Medical history	
Cardiovascular comorbidities %(n) Atrial fibrillation Hypertension	Categorical
Comorbidities %(n) COPD Diabetes CKD Liver disease	Categorical
Smoking (former or active) %(n)	Categorical
Alcohol over 7 /14 units pr. week (woman/men) %(n)	Categorical
Laboratory values	
Haemoglobin mmol/L	Continuous
Sodium mmol/L	Continuous
Potassium mmol/L	Continuous
Creatinine micromole/L	Continuous
eGFR (mL/min/1.73m2)	Continuous
NT-proBNP (pmol/L)	Continuous
Albumin g/L	Continuous

Troponin T/I above the upper reference limit %(n)	Categorical
C reactive protein mg/L	Continuous
Treatment	
ACEi/ARB/ARNI %(n)	Categorical
Beta-blockers %(n)	Categorical
MRA %(n)	Categorical
SGLT2i %(n)	Categorical
Home loop diuretics %(n)	Categorical
Home loop diuretics dose (furosemide-equivalent dose, mg daily)	Continuous
Anticoagulant therapy %(n) VKA DOAC LMWH	Categorical
ICD %(n)	Categorical
CRT %(n)	Categorical

Statistical method for comparison of baseline characteristics

For descriptive statistics, normally distributed continuous variables will be reported as mean ± standard deviation (SD), skewed distributed variables as median and interquartile range (IQR) and categorical variables as numbers (n) and percentages (%).

Outcome definitions and timing of outcome assessment

A detailed standard operating procedure (SOP ver. 2.0 TAP-IT data quality, collection, and definition version 28.07.22) for data collection has been prepared (key definitions summarised below)

Outcome	Type	Assessment time
Primary outcome		
Number of days alive outside hospital during the 90 days following randomisation. Definition: A readmission is defined as admission to any department in the hospital due to any cause lasting over 24 hours or a change in calendar date.	Count, days	90 days after randomisation.
Secondary outcomes		

Number of days alive and not hospitalised due to heart failure during the 90 days following randomisation. Definition: Readmissions meeting the above-mentioned criteria and related to heart failure based on the definition provided by The American Heart Association [11]	Count, days	90 days after randomisation.
Duration of index admission Definition: The start date for hospitalisation in any department. The Department is registered, and the patients can be transferred between departments and hospitals (e.g. to specialised heart centres) during admission. The end of admission is defined as when the patient is discharged, either to their home facility or to care in the primary sector (nursing home, rehabilitation centre)	Count, days	From randomisation until the date of discharge from the hospital.
Changes from baseline in weight during admission - Weight obtained within 24 hours from randomisation and discharge.	Continuous, kg	From randomisation until the date of discharge from the hospital.
Change from baseline in dosage of diuretics during admission - Daily dose at discharge minus daily dose at baseline (both within 24 hours) - Maximum daily dose during admission minus daily dose at baseline (within 24 hours)	Furosemide-equivalent dose, Mg	From randomisation until the date of discharge from the hospital.
Time to death	Time to event, Days	90 days after randomisation.
Time to first readmission or death Definition: A readmission is defined as admission to any department in the hospital due to any cause lasting over 24 hours or a change in calendar date.	Time to event, Day	90 days after randomisation.

Number of complications to interventional thoracentesis: - Pneumothorax. - Intrapleural bleeding (haemothorax) - Infection requiring antibiotics (empyema) - Pain/discomfort with the need for analgesics - Organ laceration - Re-expansion pulmonary oedema. - Thromboembolic events within 30 days from discontinued anticoagulation therapy (deep vein thrombosis, pulmonary embolism, ischemic stroke, transient ischemic attack, systemic embolism, myocardial infarction)	Count	From randomisation until the date of discharge from the hospital.
Number of complications during the index admission - Delirium requiring pharmacological therapy - Fall - Nosocomial infections like pneumonia, urine catheter-related infection, and central venous catheter-related infections requiring a change of the catheter and antibiotic treatment.	Count	From randomisation until the date of discharge from the hospital.
Kansas City Cardiomyopathy Questionnaire score (KCCQ-23)	Scores from 0-100. 100 representing the best outcome.	14 days after discharge from index admission and 90 days after randomisation
Satisfaction with hospital stay - Selected questions from the questionnaire "Questions about your admission" from the annual Danish National Survey of Patient Experiences”	Likert scale from 1-5. 5 represents the best outcome.	14 days after discharge from index admission

Primary outcome definition

Number of days alive outside of the hospital during the 90 days following randomisation. The outcome is countable with a minimum value of 0 and a maximum of 90. In case of death before the end of follow-up, the participant will be assigned the number of days they were outside of the hospital until death occurred. If the participant dies during the index admission, they will be assigned 0 days alive and outside of the hospital (Auriemma et al., 2021; Granholm et al., 2023). We anticipate 95% of participants to be discharged from the hospital [5–7] and based on limited previous observation data in patients with heart failure-related pleural effusion necessitating thoracentesis, we anticipated 78-91% of participants to survive 30 days after hospitalisation [1,2].

Missing data for the primary outcome

Complete follow-up of all enrolled participants will be attempted. Primary endpoint data regarding death and admission days will be collected from the participant's electronic medical record regardless of whether questionnaire follow-up was completed. Therefore, we anticipate complete data for the primary outcome since the only way that data will not be available is due to immigration within the 90-day follow-up and investigators not being able to get in contact with the participant. The extent of missing primary outcome data will be reported. The main analysis will be a complete case analysis on data as observed. In case of missingness for the primary endpoint, sensitivity analysis will be conducted for the best-worst and the worst-best case scenario. For this purpose, we define a “harmful outcome” (worst) as no further days alive since the participant was lost to follow-up. A “beneficial outcome” (best) will be that the participant was alive and out of the hospital for the rest of the period from they were lost to follow-up till the end of the follow-up.

Missing data for the secondary outcome

The extent of missing outcome data will be reported and discussed. Missing data from questionnaires on quality of life and patient satisfaction are anticipated and, in some cases due to the participants being deceased. The number and percentages of participants that completed questionnaire follow-up at 14 days and 90 days will be reported. For participants who did not complete the questionnaire during follow-up, an explanation will be provided if available. If an entire questionnaire is missing, we assume that this could be due to deteriorating health or death and the data is therefore missing not at random (MNAR) [15]. The main analysis will be as observed, and sensitivity analysis will be conducted for the best-worst and the worst-best case scenario. Patients who have died will be assigned the worst score (0), and patients who are alive and have not responded despite efforts to collect data (electronic reminders and telephone contact) are assumed to be of very poor to poor health and assigned the worst score (0). Missing data from

questionnaires on satisfaction with hospital stay (Likert scale 1-5, 5 representing highest satisfaction) follows the same assumption as KCCQ with data MNAR. However, it is not assumed that patients with poor health status or who die during follow-up necessary were dissatisfied with their treatment. Therefore, simple mean imputation will be used.

Statistical methods to compare primary and secondary outcomes between groups

Main analysis of primary outcome

The primary outcome of the trial, "days alive without hospitalisation during the 90 days following randomisation", is a countable outcome with an anticipated non-normal distribution.
Accordingly, we plan a non-parametric analysis of the primary outcome [16]. We will assess the Mann-Whitney parameter using the Wilcoxon-Mann-Whitney-test on the full analysis dataset according to the intention-to-treat principle [17].
For explorative purposes, the analysis will also be performed on the per-protocol dataset and as-treated dataset.

Sensitivity analyses

In addition to the sensitivity analyses described to assess the impact of missing data. The following sensitivity analyses for the primary outcome will be performed and reported to assess the robustness of the result of the main analysis in the intention-to-treat population.

- the van Elteren test with no strata
- the van Elteren test stratified by site
- the van Elteren test stratified by anticoagulation status.

Analyses of Secondary Outcomes

The number of days alive and not hospitalised due to heart failure has the same assumptions as the primary endpoint and will be analysed as such. Distribution of the secondary outcome continues variables will be determined by visual inspection of the histogram and by QQ plots. Normally distributed continuous variables will be compared by students' t-test and non-normal distributed continuous variables by the non-parametric Wilcoxon-Mann-Whitney-test. Categorical data will be compared using Chi-squared or exact Fisher test if n<5 in a group. Time-to-event data will be compared by the log-rank method, and adjusted analyses will be performed by proportional hazards regression. The main analysis of KCCQ scores will be a comparison of the mean score at each time point. For further exploratory analyses, the total KCCQ-score at each time point will be summarized in 25-point ranges, where scores represent health status as follows: 0 to 24: very poor to poor; 25 to 49: poor to fair; 50 to 74: fair to good; and 75 to 100: good to excellent [19].

The number and percentage of complications to thoracentesis will be reported. The number and percentage of thromboembolic events within 30 days and complications during the index admission will be compared between the groups. Results will be presented in tables, and time-to-event outcomes will be illustrated by Kaplan-Meier curves.

Subgroup analyses

The following prespecified subgroups will be analysed to evaluate the heterogeneity of any treatment effect.

- Age above/below median age
- Sex (male vs. female)
- Anticoagulation therapy (yes/no)
- Albumin level (over and under mean level)
- LVEF (over and under mean level)
- Severe chronic kidney disease (eGFR < 30) (yes/no)

Harms

Both treatment strategies are established practices with known side effects and complications. The study is not powered to investigate the comparative effectiveness of the intervention and comparator regarding safety and harm. The number of complications to interventional thoracentesis in the intervention group and complications during the index admission in both groups will be assessed as secondary outcomes, as specified in Table 2.

Statistical software:

All analysis will be performed in the statistical software R in the release version available at the end of the follow-up.

References

1 Debiasi EM, Pisani MA, Murphy TE, *et al.* Mortality among patients with pleural effusion undergoing thoracentesis. *European Respiratory Journal*. 2015;46:495–502.

2 Walker SP, Morley AJ, Stadon L, *et al.* Nonmalignant Pleural Effusions: A Prospective Study of 356 Consecutive Unselected Patients. *Chest*. 2017;151:1099–105.

3 McDonagh TA, Metra M, Adamo M, *et al.* 2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure. *Eur Heart J*. 2021;00:1–128.

4 Heidenreich PA, Bozkurt B, Aguilar D, *et al.* 2022 AHA/ACC/HFSA Guideline for the Management of Heart Failure: a report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. *J Am Coll Cardiol*. 2022;xx:xxx–xxx.

5 Lassus JPE, Siirilä-Waris K, Nieminen MS, *et al.* Long-term survival after hospitalization for acute heart failure - Differences in prognosis of acutely decompensated chronic and new-onset acute heart failure. *Int J Cardiol*. 2013;168:458–62.

6 Nieminen MS, Brutsaert D, Dickstein K, *et al.* EuroHeart Failure Survey II (EHFS II): A survey on hospitalized acute heart failure patients: Description of population. *Eur Heart J*. 2006;27:2725–36.

7 Kapoor JR, Kapoor R, Ju C, *et al.* Precipitating Clinical Factors, Heart Failure Characterization, and Outcomes in Patients Hospitalized With Heart Failure With Reduced, Borderline, and Preserved Ejection Fraction. *JACC Heart Fail*. 2016;4:464–72.

8 Yelland LN, Sullivan TR, Voysey M, *et al.* Applying the intention-to-treat principle in practice: Guidance on handling randomisation errors. *Clinical Trials*. 2015;12:418–23.

9 Rockwood K, Song X, MacKnight C, *et al.* A global clinical measure of fitness and frailty in elderly people. *CMAJ*. 2005;173:489–95.

10 Sze S, Zhang J, Pellicori P, *et al.* Prognostic value of simple frailty and malnutrition screening tools in patients with acute heart failure due to left ventricular systolic dysfunction. *Clinical Research in Cardiology*. 2017;106:533–41.

11 Hicks KA, Mahaffey KW, Mehran R, *et al.* 2017 Cardiovascular and Stroke Endpoint Definitions for Clinical Trials. *Journal of the American College of Cardiology* . 2018;71:1021–34.

12 Granholm A, Kaas-Hansen BS, Lange T, *et al.* Use of days alive without life support and similar count outcomes in randomised clinical trials – an overview and comparison of methodological choices and analysis methods. *BMC Med Res Methodol*. 2023;23. doi: 10.1186/s12874-023-01963-z

13 Auriemma CL, Taylor SP, Harhay MO, *et al.* Hospital-free days: A pragmatic and patient-centered outcome for trials among critically and seriously Ill patients. *Am J Respir Crit Care Med*. 2021;204:902–9.

14 Jakobsen JC, Gluud C, Wetterslev J, *et al.* When and how should multiple imputation be used for handling missing data in randomised clinical trials - A practical guide with flowcharts. *BMC Med Res Methodol*. 2017;17:1–10.

15 Carreras G, Miccinesi G, Wilcock A, *et al.* Missing not at random in end of life care studies: multiple imputation and sensitivity analysis on data from the ACTION study. *BMC Med Res Methodol*. 2021;21. doi: 10.1186/s12874-020-01180-y

16 Jakobsen JC, Tamborrino M, Winkel P, *et al.* Count Data Analysis in Randomised Clinical Trials. *J Biom Biostat*. 2015;6:227.

17 Fay MP, Malinovsky Y. Confidence intervals of the Mann-Whitney parameter that are compatible with the Wilcoxon-Mann-Whitney test. *Stat Med*. 2018;37:3991–4006.

18 Sawilowsky SS. Misconceptions leading to choosing the t test over the wilcoxon mann-whitney test for shift in location parameter. *Journal of Modern Applied Statistical Methods*. 2005;4:598–600.

19 Spertus JA, Jones PG, Sandhu AT, *et al.* Interpreting the Kansas City Cardiomyopathy Questionnaire in Clinical Trials and Clinical Care. *J Am Coll Cardiol*. 2020;76:2379–90.