

MASTEROPPGAVE
Master i fysioterapi
Mai 2019

Samsvar mellom selvrapportert og objektivt målt fysisk aktivitet blant eldre ryggpasienter – en validitets- og feasibilitystudie av IPAQ-SF og AX3 akselerometer

Emil Eirik Kvernberg Thomassen



OsloMet – storbyuniversitetet

Fakultet for helsevitenskap
Institutt for fysioterapi

Forord

Først og fremst vil jeg takke veilederne Kjersti Storheim og Dagfinn Matre for utmerket støtte og veiledning i dette prosjektet. En stor takk rettes også til BACE-N ved professor Margreth Grotle, stipendiater Rikke Munk Killingmo og Ørjan Vigdal for godt samarbeid underveis i dette masterprosjektet. Sist, men ikke minst, rettes stor takknemlighet til Statens Arbeidsmiljøinstitutt (STAMI) for kontorplass og hjelp til prosessering av akselerometer-data.

Sammendrag

Bakgrunn: Ryggsmerter er en ledende årsak til funksjonsbegrensninger og særskilt hos eldre (>55 år) er ryggsmerter en vanlig helseplage. Fysisk aktivitet er anbefalt både som forebygging og behandling av ryggsmerter. Det har imidlertid vist seg utfordrende å måle fysisk aktivitet og det er derfor nødvendig med måleinstrumenter som er anvendelige og av god metodisk kvalitet dersom helsepersonell skal anbefale og evaluere fysisk aktivitet.

Hensikt: Å undersøke samtidig validitet til International Physical Activity Questionnaire Short Form (IPAQ-SF) med AX3 akselerometer som kriteriestandard i et utvalg eldre voksne som har oppsøkt primærhelsetjenesten for ryggsmerter. Feasibility på både IPAQ-SF og AX3 akselerometer ble også undersøkt.

Metode og materiale: 10 deltakere fra deltakerbasen til BACE-N (Back complaints in elders-Norway) fullførte IPAQ-SF og gikk med AX3 akselerometer i totalt 7 døgn. To ulike prosesseringsmetoder (Metode1_{Esliger} og Metode2_{Doherty}) på rådata fra AX3 akselerometer. Spearman rang-analyse ble benyttet for undersøkelse av validiteten til IPAQ-SF-intensitetene gange, moderat og meget anstrengende fysisk aktivitet. Spørreskjema med 5-punkts Likert Skala for å kartlegge feasibility ble besvart av deltakerne.

Resultater: Median alder for utvalget var 65 år, median smerte (Numeric Rating Scale (range 0-10)) var 1.50 og median score på Roland Morris Disability Questionnaire (range 0-24) var 0. Målinger av gange viste korrelasjon på 0.264 mellom IPAQ-SF og Metode2_{Doherty}. Ved måling av moderat fysisk aktivitet med IPAQ-SF og Metode2_{Doherty} var korrelasjonen 0.733 ($p=0.025$). Negativ korrelasjon på -0.233 ble funnet mellom Metode1_{Esliger} og IPAQ-SF ved måling av moderat aktivitet. Måling av meget anstrengende aktivitet med IPAQ-SF viste ingen korrelasjon med Metode1_{Esliger}. To IPAQ-SF ble returnert ukomplett, mens måling med akselerometer produserte median 6.94-6.98 døgn med måledata. Vurdering av feasibility viste at deltakerne fant både IPAQ-SF og AX3 akselerometer å være akseptable målemetoder. Avhengig av målemetode (IPAQ-SF, Metode1_{Esliger} eller Metode2_{Doherty}) var henholdsvis 70, 100 og 90% av deltakerne å regne som tilfredsstillende fysisk aktive. Korrelasjon mellom Metode1_{Esliger} og Metode2_{Doherty} var 0.576 og 0.462 på henholdsvis lett og moderat fysisk aktivitet.

Konklusjon: IPAQ-SF hadde tilfredsstillende samtidig validitet (0.733) med AX3 akselerometer ved måling av moderat fysisk aktivitet når Metode2_{Doherty} benyttes, mens akselerometerdata prosessert med Metode1_{Esliger} ga motstridende resultat med hensyn til samtidig validitet (-0.233 (p=0.546), hvilket tyder på en diskrepans mellom Metode1_{Esliger} og Metode2_{Doherty}. Måling av gange med IPAQ-SF viste svak samtidig validitet (0.264) med Metode2_{Doherty}. Både IPAQ-SF og AX3 akselerometer var regnet som akseptable målemetoder, men med liten fordel i favør AX3. Grunnet lavt utvalg må resultatene tolkes med forsiktighet.

Nøkkelord: IPAQ-SF, AX3 akselerometer, Fysisk aktivitet, Ryggsmerter, Eldre

Summary

Background: Back pain is a leading cause of disability, more prominent within the elderly population (>55 years) is back pain a common health issue. Physical activity is recommended for prevention and treatment of back pain. However, there are challenges associated with measuring activity and applicable measurement tools with good methodological quality are a necessity for healthcare workers when recommending and evaluating physical activity.

Purpose: To assess the concurrent validity of the International Physical Activity Questionnaire Short Form (IPAQ-SF) with the AX3 accelerometer as criteria standard in a selection of elderly seeking primary healthcare due to back pain. Feasibility on both IPAQ-SF and AX3 accelerometer was also assessed.

Method and material: 10 participants from the BACE-N (Back Complaints in elders-Norway) pool completed the IPAQ-SF and wore the AX3 accelerometer for a total of 7 days. Two different processing methods (Method1_{Esliger} and Method2_{Doherty}) were utilized when processing raw accelerometer data from AX3. Spearman rank-correlation was determined for calculating the validity IPAQ-SF intensity levels walking, moderate and vigorous physical activity. A questionnaire with 5-points Likert scale for evaluating feasibility on IPAQ-SF and AX3 were completed.

Results: Median age for the selection was 65, median pain (Numeric Rating Scale (range 0-10)) was 1.50 and median score for Roland Morris Disability Questionnaire (range 0-24) was 0. When measuring walking correlation showed 0.264 between IPAQ-SF and Method2_{Doherty}. When measuring moderate physical activity with IPAQ-SF and Method2_{Doherty} analysis showed a correlation coefficient of 0.733 ($p=0.025$). Negative correlation (-0.233) was found between IPAQ-SF and Method1_{Esliger} when measuring moderate activity. Measuring vigorous activity with IPAQ-SF showed no correlation with Method1_{Esliger}. Two IPAQ-SF questionnaires was returned incomplete, while AX3 produced median 6.94-6.98 days of valid measurement data. Feasibility of IPAQ-SF and AX3 showed that participants found both methods acceptable. Depending on measurement methods (IPAQ-SF, Method1_{Esliger} and Method2_{Doherty}) 70, 100 or 90% of the participants were satisfactory physical active. Correlation between Method1_{Esliger} and Method2_{Doherty} showed correlation coefficients of 0.576 and 0.462 respectively when measuring light and moderate physical activity.

Conclusion: IPAQ-SF showed satisfactory concurrent validity with AX3 accelerometer in measuring moderate physical activity when Method2_{Doherty} is utilized, while utilizing Method1_{Esliger} showed conflicting results regarding concurrent validity (-0.233). Which implies that there is a discrepancy between Method1_{Esliger} and Method2_{Doherty}. Measuring walking showed low correlation between IPAQ-SF and Method2_{Doherty}. Both IPAQ-SF and AX3 were considered as acceptable methods of measurement, with a slight favour for AX3. Due to a small selection, results must be interpreted carefully.

Keywords: IPAQ-SF, AX3 accelerometer, Physical activity, Back pain, elderly

Innholdsfortegnelse

DEFINISJONER/FORKORTELSER/BEGREPSAVKLARINGER:	3
1.0 INTRODUKSJON	4
1.1 FORMÅL OG HENSIKT	5
1.2 PROBLEMSTILLING, HYPOTESE OG FORSKNINGSSPØRSMÅL	5
2.0 TEORETISK BAKGRUNN	6
2.1 RYGGSMERTER	6
2.1.1 Forebygging, tiltak og behandling ved ryggsmertesmerter	8
2.1.2 Ryggsmertesmerter og eldre	10
2.1.3 Ryggsmertesmerter og fysisk aktivitet	11
2.2 FYSISK AKTIVITET	12
2.2.1 Måling av fysisk aktivitet	13
2.2.2 Subjektiv eller selvrapportert måling av fysisk aktivitet	15
2.2.3 IPAQ-SF	16
2.2.4 Objektiv måling av fysisk aktivitet	18
2.2.5 AX3 Akselerometer	20
2.3 MÅLEEGENSKAPER	23
2.3.1 Validitet	23
2.3.2 Feasibility	24
3.0 METODE	26
3.1 DESIGN	26
3.2 REKRUTTERING OG UTVALG	26
3.3 ETIKK	28
3.4 VARIABLER	28
3.4.1 Deskriptive variabler	28
3.4.2 International Physical Activity Questionnaire (IPAQ-SF)	29
3.4.3 Axivity AX3 akselerometer	30
3.4.4 Metode1 ^{Esliger} Wear-time validering, kvantifisering og klassifisering av rådata til aktivitetsdata	31
3.4.5 Metode2 ^{Doherty} Wear time validering, kvantifisering og klassifisering av rådata til aktivitetsdata	33
3.5 FEASIBILITY	34
3.6 STATISTIKK	35
3.6.1 Deskriptiv statistikk	35
3.6.2 Samtidig validitet	36
3.6.3 Feasibility	36
4.0 RESULTATER	37
4.1 UTVALG	37
4.2 AKTIVITETSDATA FRA IPAQ-SF OG AX3	40
4.3 SAMTIDIG VALIDITET	40
4.4 KORRELASJON METODE1 ^{ESLIGER} OG METODE2 ^{DOHERTY}	41
4.5 FEASIBILITY	41
5.0 DISKUSJON	43
5.1 RESULTATER	43
5.1.1 Samtidig validitet	43
5.1.2 Korrelasjon mellom Metode1 ^{Esliger} og Metode2 ^{Doherty}	46
5.1.3 Feasibility	47
5.1.4 Statistiske valg	48
5.2 METODE	49
5.2.1 Design	49
5.2.2 Rekruttering og utvalg	50
5.2.3 Etikk	54
5.2.4 Aktivitetsdata fra rapportering med IPAQ-SF	54
5.2.5 Aktivitetsdata fra måling med Axivity AX3	56

5.2.6 Metode1Esliger og Metode2Doherty	57
5.2.7 Feasibility.....	61
5.3 KLINISKE IMPLIKASJONER / VIDERE FORSKNING	61
6.0 KONKLUSJON	63
7.0 APPENDIKS	65
8.0 REFERANSER	72

Definisjoner/forkortelser/begrepsavklaringer:

Activity AX3 Akselerometer: Tri-aksialt akselerometer som benyttes som objektiv metode i denne oppgaven.

BACE-N: Back complaints in elders-Norway. Representerer den norske grenen av en større internasjonal multisenter-studie på ryggsmarter blant eldre.

Feasibility: Begrep som omhandler gjennomførbarheten til intervensjoner og måleverktøy som benyttes i studier (Bowen et al., 2009).

Fysisk aktivitet: Enhver kroppslig bevegelse produsert av skjelettmuskler som resulterer i energiforbruk (Caspersen, Powell, & Christenson, 1985).

Intensitetsnivå: Beskrivelse av intensiteten til hvilken som helst aktivitet, eksempelvis lett, moderat eller anstrengende fysisk aktivitet.

IPAQ-SF: International Physical Activity Questionnaire Short Form

MET: Metabolic equivalent task. Mål på energiforbruk i ulike typer aktivitet.

Modifiserbare faktorer / ikke-modifiserbare faktorer: Knyttet til Cognitive Functional Therapy og er begreper som benyttes for å beskrive ulike faktorer som kan være avgjørende for graden av ryggsmarter. Med modifiserbare faktorer menes eksempelvis fysisk aktivitet, mens ikke-modifiserbare faktorer kan eksempelvis være alder, kjønn eller barndom (O'Sullivan et al., 2018).

Objektiv målt fysisk aktivitet: Med objektivt målt fysisk aktivitet i denne oppgaven menes aktivitet målt med akselerometer.

Prosesseringsmetode: Begrep som beskriver metode for å kvantifisere og klassifisere rådata fra akselerometer

Subjektiv målt fysisk aktivitet: Med subjektiv målt fysisk aktivitet i denne oppgaven menes aktivitet målt med selvrappoterer gjennom IPAQ-SF

Wear-/nonwear time: Begrep som brukes om hvorvidt deltakere har hatt på seg akselerometeret eller ikke.

1.0 Introduksjon

Ryggsmerter er en ledende årsak til funksjonsbegrensninger (Vos et al., 2017). Særskilt blant eldre er ryggsmerter en vanlig helseplage og blant amerikanere i aldersgruppen 45-74 år er punktprevalensen 28-29% (Deyo, Mirza, & Martin, 2006; Wong, Karppinen, & Samartzis, 2017). Tall fra Norge viser at ryggsmerter er en hyppig årsak til kontakt med helsevesenet og punktprevalensen viser at henholdsvis 23.7 % og 30.3 % av menn og kvinner i aldersgruppen 60-79 år rapporterer ryggsmerter (I. Heuch, Hagen, Heuch, Nygaard, & Zwart, 2010; Kinge, Knudsen, Skirbekk, & Vollset, 2015). Til tross for høy prevalens er eldre en underrepresentert gruppe i forskningen på ryggsmerter, noe som gir grunn til å se nærmere på feltet (Wong et al., 2017).

En rekke ulike behandlingstiltak som massasje, mobiliseringsøvelser og kognitiv intervensjon er beskrevet i retningslinjer fra ulike land, som anbefalt behandling for pasienter med ryggsmerter. I de fleste retningslinjene er fysisk aktivitet en fellesnevner for anbefalte tiltak enten som forebyggende tiltak eller som behandling for ryggsmerter (Bernstein, Malik, Carville, & Ward, 2017; Qaseem, Wilt, McLean, Forciea, & Clinical Guidelines Committee of the American College of, 2017). En oversiktsartikkel viste at fysisk aktivitet på fritiden kan redusere risikoen for kroniske ryggsmerter med 11-16 % (Shiri & Falah-Hassani, 2017). Det er også vist at forskjellige former for fysisk aktivitet som eksempelvis aerobic, thai-chi, styrketrening og yoga reduserer smerte ved en rekke kroniske smertetilstander, inkludert ryggsmerter (Geneen et al., 2017).

Ifølge norske målinger gjennomført i tidsrommet 2014 og 2015, er det generelt lavere nivåer av fysisk aktivitet hos individer over 65 år, med et betydelig fall etter fylte 75 år (Helsedirektoratet, 2015, s. 7). Et økende antall eldre i Norge gjør at det er viktig med gode og validerte målemetoder for å kartlegge fysisk aktivitet i den eldre delen av befolkningen. Måling av fysisk aktivitet, uavhengig av om man bruker subjektive eller objektive målemetoder, har imidlertid vist seg å være utfordrende hos befolkningen og ikke minst eldre (Kowalski, Rhodes, Naylor, Tuokko, & MacDonald, 2012; Sanda et al., 2017). Dersom helsepersonell skal kunne anbefale fysisk aktivitet som forebygging og behandling er det viktig med målemetoder som er anvendelige og av god metodisk kvalitet for evaluering av aktivitetsnivået til pasienter (Falck, McDonald, Beets, Brazendale, & Liu-Ambrose, 2016).

1.1 Formål og hensikt

BACE-N-studien (Back Pain in Elderly-Norway) er en del av en internasjonal kohortestudie som kartlegger ryggsmerte, funksjonsnivå, behandling og prognostiske faktorer hos eldre ryggpasienter (>55 år) i primærhelsetjenesten (Scheele et al., 2011). I studien benyttes International Physical Activity Questionnaire Short Form (IPAQ-SF) som mål på fysisk aktivitet. IPAQ-SF er et 9-items spørreskjema for kartlegging av siste ukes fysiske aktivitet og blir benyttet på samtlige oppfølgingstidspunkt (baseline, 3, 6, 12 og 25 måneder) i BACE-N. Det er derfor viktig å ha kunnskap om skjemaets psykometriske egenskaper. Målet med dette masterprosjektet er å undersøke den samtidige validiteten til den norske versjonen av spørreskjemaet IPAQ-SF-skjemaet anvendt på ryggpasienter >55 år ved å validere det med en objektiv målemetode i form av akselerometer. I tillegg undersøkes feasibility av IPAQ-SF og AX3 akselerometer.

1.2 Problemstilling, hypotese og forskningsspørsmål

Dette masterprosjektet tar for seg den samtidige validitet til IPAQ-SF målt med AX3 akselerometer, samt feasibility til disse to målemetodene gjennom problemstillingen: *«Er det samsvar mellom selvrappportert fysisk aktivitet registrert ved bruk av IPAQ-SF, og objektiv målt fysisk aktivitet registrert ved bruk av AX3 hos eldre pasienter med ryggsmerte?»*

Nullhypotesen (H0) for denne masteroppgaven er: *«Det er ikke forskjell mellom IPAQ-SF og AX3 akselerometer i registrering av fysisk aktivitet.»*

Forskningsspørsmål:

Hvordan er den samtidige validiteten til IPAQ-SF hos pasienter >55 år med ryggsmerte?

Hvordan er feasibility til henholdsvis IPAQ-SF og AX3 akselerometer?

Masterprosjektet ønsker i tillegg å eksplorere følgende forhold:

- Møter utvalget som er undersøkt kravene for fysisk aktivitet (150 minutter moderat fysisk aktivitet/75 minutter anstrengende fysisk aktivitet per uke)?
- Har valg av prosesseringsmetode på akselerometerdata innvirkning på antall minutter fysisk aktivitet hos deltakerne?

2.0 Teoretisk bakgrunn

I forkant av datainnsamling ble det gjort en rekke litteratursøk på feltet for å kartlegge hva som fantes av lignende valideringsstudier av IPAQ-SF. Et større systematisk søk ble gjennomført med søkeordene [*international physical activity questionnaire*] (and) [*validity*]. Søkene ble gjennomført sommeren 2018, og ble gjentatt regelmessig utover høsten og vinteren 2018/2019 for å fange opp eventuelle nyere studier som var av interesse. Det ble også fanget opp aktuelle studier gjennom handsøk fra litteraturlistene til allerede inkluderte studier. Disse studiene er anvendt i det teoretiske grunnlaget for denne oppgaven.

Majoriteten av den publiserte litteraturen på feltet ryggsmert er knyttet til lave ryggsmert som anatomisk strekker seg fra første sakrale virvel til første lumbalvirvel. BACE-N-studien inkluderer pasienter med smert fra første sakrale virvel til toppen av skulderbladene (Scheele et al., 2011). Det betyr altså et noe utvidet anatomisk område i ryggen. Litteraturen er imidlertid i stor grad sammenfallende og skiller i liten grad mellom lumbale og thoracale ryggsmert. Et eksempel på dette er en nylig publisert studie fra BACE gjennomført i Brasil av F. R. D. Jesus-Moraleida et al. (2018) hvor definisjonene nevnt ovenfor benyttes likt.

2.1 Ryggsmert

Som innledning til den nåværende forståelsen av ryggsmert er det hensiktsmessig med et kort historisk tilbakeblikk på hvordan man har forsket på og forstått ryggsmert. Det har langt tilbake i tid eksistert dokumentasjon på ryggsmert blant mennesker (Allan & Waddell, 1989). På begynnelsen av 1900-tallet eksisterte det flere teorier om årsaken til ryggsmert. To eksempler på forklaringene av ryggsmert var at de enten kom fra nevropati eller revmatiske lidelser (Lutz, Butzlaff, & Schultz-Venrath, 2003). Innføringen av radiologisk undersøkelse gjorde at man hos flere pasienter oppdaget uregelmessigheter i mellomvirvelskivene. Denne oppdagelsen førte videre til at man fra 1938 og utover store deler av 1900-tallet vektla mellomvirvelskivene som årsak til ryggsmert (Lutz et al., 2003). Dette er i strid med dagens kunnskap hvor radiologisk undersøkelse ikke er anbefalt som del av en standardisert undersøkelse, med mindre det foreligger mistanke om alvorlig underliggende patologi (Foster N, 2018).

Ryggsmerter defineres vanligvis som smerter fra nederste ribbe til begynnelsen av setet (Dionne et al., 2008). Ryggsmertenes varighet betegnes som: akutt (<4 uker), sub-akutt (4-12 uker) og kronisk ryggsmerte (>12 uker) (Qaseem et al., 2017). I den nåværende forståelsen av ryggsmerter skiller man også mellom spesifikke og uspesifikke ryggsmerter, korsryggsmerter med nerverotsaffeksjon (prolaps og spinal stenose) og ryggsmerter som skyldes underliggende patologi som for eksempel infeksjoner, betennelsestilstander eller brudd (Laerum, Storheim, & Brox, 2007). I vestlige land er prevalensen av uspesifikke ryggsmerter, ryggsmerter som skyldes nerverotsaffeksjon og underliggende patologi henholdsvis 80-90 %, 5-10 og 1-5% (Laerum et al., 2007). Disse tallene viser at alvorlige patologiske årsaker sjeldent ses hos pasienter med ryggsmerter, og at man for flertallet av ryggpasientene ikke makter å identifisere en spesifikk nociseptisk årsaksforklaring (Hartvigsen et al., 2018). Pasientene med mangel på klar patologisk/nociseptisk årsaksforklaring blir derfor kategorisert og behandlet under begrepet «uspesifikke ryggsmerter». Uspesifikke ryggsmerter blir gjerne forklart ut i fra den biopsykososiale modellen (Maher, Underwood, & Buchbinder, 2017).

Den biopsykososiale modellen ble presentert av Georg Engel i 1977 som et motsvar til den biomedisinske tenkningen (Engel, 1977). I etterkant av dette har den biopsykososiale modellen fungert som et rammeverk for uspesifikke ryggsmerter hvor både biologiske faktorer, psykologiske og sosiale faktorer antas å utgjøre en viktig del av forståelsen. Dette har ført til et skifte i behandling fra biomedisinsk behandling, som kirurgi og farmakologi, til kognitive- og selvmestringsterapier (Foster et al., 2018). Forståelsen av ryggsmerter har altså endret seg gjennom tidene fra teorier om ryggsmerter som følge av traumer i anatomisk struktur til den nåværende forståelsen av ryggsmerter som et resultat av flere biopsykososiale faktorer (Allan & Waddell, 1989; Hartvigsen et al., 2018).

I en omfattende rapport fra World Health Organization (WHO) publisert i LANCET har man estimert prevalens, insidens og antall år med nedsatt funksjon på 328 sykdommer i 192 ulike land. Rapporten av Vos et al. (2017) viser at ryggsmerter er den hyppigste enkeltårsak til at mennesker rapporterer funksjonsbegrensninger (years lived with disability). I tillegg er det fra 1990-tallet og frem til 2015 rapportert en økning på 54% i antall år levd med nedsatt funksjon grunnet ryggsmerter ("Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015," 2016; Vos et al., 2017). Økningen i individer med ryggsmerter er også lik på tvers av lav-, midt- og høy-inntektsland – noe som skiller ryggsmerter fra en del

andre lidelser (Vos et al., 2017). Tallene på muskel og skjelettlidelser i rapporten av Vos et al. (2017) er hentet fra selvrapporteringskjema. En av flere årsaksforklaringer på økningen i antall dager med funksjonsbegrensninger på grunn av rygg smerter er det økende antall eldre (Hoy et al., 2012; Hoy et al., 2010; Vos et al., 2017). På nasjonalt nivå er det lignende tendenser hvor rygg- og nakkelidelser er den største årsaken til sykefravær blant arbeidstagere med muskel- og skjelettlidelser (Brage, Ihlebaek, Natvig, & Bruusgaard, 2010). Videre viser tall fra Norge at omtrent 80 % vil oppleve smerter i muskel- og skjelettsystemet i løpet av en måned (Ihlebaek, Brage, Natvig, & Bruusgaard, 2010).

2.1.1 Forebygging, tiltak og behandling ved rygg smerter

Ved forebygging av rygg smerter er det hensiktsmessig å dele inn i primær og sekundær forebygging. Primær forebygging handler om tiltak som retter seg mot å minske insidensen, altså nye tilfeller rygg smerter, mens sekundær forebygging baserer seg på å forhindre videre komplikasjoner av rygg smerter (Tulchinsky, Varavikova, & Varavikova, 2000, p. 66). En systematisk oversiktsartikkel viste at trening og undervisning betegnes som moderat effektive tiltak for primær forebygging av rygg smerter (Steffens et al., 2016). For sekundær forebygging viste en systematisk oversiktsartikkel av 36 prospektive kohort-studier at man kan redusere risikoen for langvarige rygg smerter med 11-16 % gjennom fysisk aktivitet på fritiden (Shiri & Falah-Hassani, 2017). Forfatterne knytter imidlertid stor usikkerhet til resultatene som fremkom i den systematiske gjennomgangen grunnet lav metodisk kvalitet til de inkluderte studiene (Shiri & Falah-Hassani, 2017). Det er videre argumentert for at man burde sette invalidiserende rygg smerter på politisk dagsorden, samt fremme endringer i helsevesenet med fokus på skifte fra et fragmentert biomedisinsk fokus til å fremme egenmestring av rygg smerter (Buchbinder et al., 2018).

For behandling av rygg smerter har man ulike behandlingstiltak, hvor det i denne oppgaven fokuseres på konservativ behandling i form av ikke-invasive behandlingstiltak. Per dags dato er den konservative formen for behandling regnet som anbefalt behandling til rygg pasienter (Shipton, 2018). Dette er grunnlaget for behandling hos fysioterapeut, som baserer seg i stor grad på behandling hvor fokus er smertelindring og tilpasning av fysisk aktivitet med mål om å hindre videre utvikling av smerter og påfølgende nedsatt funksjon (Maher et al., 2017; Shipton, 2018).

De siste retningslinjene fra NICE (Britiske National Institute for Health and Care Excellence) på undersøkelse og behandling av ryggsmarter ble utgitt i 2017 (Bernstein et al., 2017). Disse retningslinjene beskriver anbefalte tiltak ved både undersøkelse, vurdering og tiltak til ryggpasienter. Undersøkelse baserer seg i hovedsak på identifisering av eventuell patologi og anvendelse av screeningverktøy for å anslå pasientenes grad av risiko for å utvikle vedvarende smerter. Felles for både høy- og lav-risiko-pasienter er tiltak for egenmestring som baserer seg på informasjon omkring smerter, samt råd om opprettholdelse av fysisk aktivitet. For høy-risiko-pasienter anbefales det intensiv oppfølging med et behandlingsopplegg som kombinerer psykologiske og fysioterapeutiske tiltak. Det vil si tiltak som manipulasjon, mobiliseringsteknikker, massasje, kognitive intervensjoner, trening og øvelser. For pasientene som vurderes å ha lavere risiko for utvikling av vedvarende smerter anbefales tiltak som gruppetrening med eksempelvis yoga, tøying og styrke (Bernstein et al., 2017).

I 2007 ble det i Norge utformet nasjonale retningslinjer for korsryggsmarter med og uten nerverotsaffeksjon (Laerum et al., 2007). Retningslinjene inneholder anbefalinger på utredning/diagnostikk og behandling/tiltak for både akutte og langvarige uspesifikke ryggsmarter og ryggsmarter med nerverotsaffeksjon. Undersøkelse og utredning er lik retningslinjene fra NICE hvor man i hovedsak skal identifisere patologi («røde flagg») og psykologiske risikofaktorer («gule flagg») som gir økt risiko for utvikling av langvarige smerter. For behandling er det på lik linje med NICE også anbefalt i norske retningslinjer at pasienter motiveres til å opprettholde normal aktivitet, både ved uspesifikke ryggsmarter og ryggsmarter med nerverotsaffeksjon. Tiltak som massasje og manipulasjon anbefales kun som supplement til tiltak på øvelser og trening. Det oppgis i retningslinjene at man ikke har dokumentasjon nok til å kunne anbefale noen primærforebyggende tiltak, men sekundærforebyggende tiltak som veiledet trening, ergonomisk tilpasning på arbeidsplass og ryggskole anbefales (Laerum et al., 2007).

I lys av den biopsykososiale modellen og forståelsen av ryggsmarter finnes det en multimodal behandling som omfatter en kombinasjon av kognitiv tilnærming og fysisk aktivitet. Cognitive functional therapy (CFT) er en behandlingsform som kombinerer komponenter fra adferdspsykologi og fysioterapi (O'Sullivan et al., 2018). Behandlingen er primært rettet mot pasientene med uspesifikke ryggsmarter. Undersøkelse og behandlingen baserer seg på å identifisere modifiserbare og ikke-modifiserbare faktorer som bidrar til den smerten som pasienten opplever. Slik skal man kunne gi pasienten en mer helhetlig forklaring på smertene

deres. Et eksempel på en modifiserbar faktor er blant annet livsstil og fysisk aktivitet. I behandling med CFT vil man derfor ha fokus på å veilede pasienten til en mer aktiv livsstil (O'Sullivan et al., 2018). En randomisert kontrollert studie på CFT konkluderte at CFT er mer effektiv som behandling sammenlignet med manuelle teknikker og øvelser (Vibe Fersum, O'Sullivan, Skouen, Smith, & Kvale, 2013). En slik type behandling kunne eksempelvis tenkes å være hensiktsmessig til pasienter som har høy risiko for å utvikle vedvarende smerter.

Oppsummert viser dette at førsteprioritet for forebygging og behandling av rygg smerter er opprettholdelse og regulering av fysisk aktivitet, samt informasjon om smerter og egenmestring ved behandling av både akutte og mer langvarige rygg smerter. Passive behandlingsformer er forbeholdt som supplement til råd om fysisk aktiv, informasjon, egenmestring, trening og øvelser.

2.1.2 Rygg smerter og eldre

Den eldre populasjonen er en underrepresentert gruppe i forskningen på rygg smerter, til tross for at dette er et stort helseproblem hos denne gruppen (Bressler, Keyes, Rochon, & Badley, 1999; Scheele et al., 2011; Weiner et al., 2003). Tallene på hvor mange eldre som rammes av rygg smerter viser en 1-års prevalens som varierer fra omtrent 11 til 50%. Tall fra Sverige viste en 1-årsprevalens blant 3009 svenske menn i aldersgruppen 69-81 hvor nær 50% hadde rygg smerter (Ghanei et al., 2014). Til sammenligning blant spanjoler i aldersgruppen 51-70 hvor man fant en 1-år prevalens på 11.45% (Palacios-Cena et al., 2015). En metaanalyse av prevalensstudier på korsrygg smerter hos brasilianske eldre viste imidlertid at blant de 28 448 undersøkte var 1-års prevalensen blant de over 60 år på 13% (Leopoldino et al., 2016). Avslutningsvis viste en systematisk oversiktsartikkel på 165 studier fra 54 land at prevalensen av rygg smerter er høyest hos aldersgruppen 40-80 og blant kvinner (Hoy et al., 2012). I samsvar med funnet fra Hoy et al. (2012) fant en prospektiv kohort-studie fra 2014 at det er en signifikant forskjell mellom utvikling av aktivitets-begrensende langvarige rygg smerter hos kvinner og menn (Makris, Fraenkel, Han, Leo-Summers, & Gill, 2014). Tidligere undersøkelser viser at kvinner generelt har mer risiko for utvikling av ulike smertetilstander, deriblant rygg smerter (Fillingim, King, Ribeiro-Dasilva, Rahim-Williams, & Riley, 2009). Dette er i samsvar med det som ble funnet i en systematisk oversiktsartikkel på den globale prevalensen av rygg smerter hvor flest andeler av kvinner rapporterte rygg smerter (Hoy et al., 2012).

2.1.3 Ryggsmarter og fysisk aktivitet

Som nevnt tidligere er fysisk aktivitet en viktig modifierbar faktor for pasienter med ryggsmarter. Denne delen av teorien belyser sammenhengen mellom grad av ryggsmarter og aktivitetsnivå, samt ulike studier gjort på fear-avoidance blant ryggpasienter og hvordan dette påvirker aktivitetsnivået.

Fear-avoidance-modellen ble først presentert i 1983 som resultat av et tverrfaglig samarbeid (Lethem, Slade, Troup, & Bentley, 1983). Modellen beskriver fear-avoidance som primært bygd på pasienters frykt for smerte og hvor unngåelse er fremtredende som håndteringsmekanisme for smerter (Lethem et al., 1983). Fear-avoidance og kinesiofobi kan sies å være to lignende begrep og brukes synonymt i litteraturen (F. R. Jesus-Moraleida et al., 2017; Knapik, Saulicz, & Gnat, 2011). Ifølge Wertli et al. (2014) korrelerer høy forekomst av fear-avoidance hos ryggpasienter med høyere smerteopplevelse og lavere funksjon, sammenlignet med ryggpasienter med lav forekomst av fear-avoidance. Denne trenden er og synlig hos pasienter med kroniske muskel- og skjelettplager hvor man ser en korrelasjon mellom høy grad av fear-avoidance og smerte (Luque-Suarez, Martinez-Calderon, & Falla, 2018). En av flere mulige årsaksforklaringer på hvorfor ryggpasienter ofte har fear-avoidance kan ligge i at mange er redde for ytterligere forverring av ryggsmertene (Stenberg, Fjellman-Wiklund, & Ahlgren, 2014).

Et viktig poeng er i hvor stor grad fear-avoidance virker inn på individers aktivitetsnivå. En studie av Carvalho et al. (2017) som målte fysisk aktivitet med akselerometer viste at fear-avoidance ikke nødvendigvis korrelerer med lavere nivå av fysisk aktivitet. Man så likevel en korrelasjon mellom frykt for bevegelse og nedsatt funksjon (Carvalho et al., 2017). En annen studie hvor man så på forskjellen i aktivitetsnivået mellom 20 friske mennesker og 29 ryggpasienter målt med akselerometer over 7 dager, fant at det var ingen signifikant forskjell mellom gruppene (van Weering, Vollenbroek-Hutten, Tonis, & Hermens, 2009). Derimot viste en tverrsnittsstudie at 86 % av pasientene på venteliste til spinal-operasjon kan defineres som fysisk inaktiv etter WHO's retningslinjer i forkant av operasjon og det ble videre funnet at høyere grad av fear-avoidance gir mindre fysisk aktivitet (Lotzke et al., 2018). Oppsummert er det derfor motstridende evidens på i hvor stor grad fear-avoidance virker inn på aktivitetsnivået til ryggpasienter.

En annen pågående diskusjon er hvorvidt det eksisterer en U-formet relasjon mellom fysisk aktivitet på fritiden og ryggsmarter. En U-formet relasjon mellom fysisk aktivitet og ryggsmarter vil si at både lavt og høyt nivå av fysisk aktivitet kan gi økt risiko for ryggsmarter (Heneweer, Vanhees, & Picavet, 2009). En tverrsnitts studie fra 2009 fant nettopp denne sammenhengen hvor man så tendenser til en U-formet relasjon mellom fysisk aktivitet og ryggsmarter. Målingene ble gjort gjennom en kohort-studie hvor spørreskjema ble benyttet som kartlegging av deltakernes fysiske aktivitet. (Heneweer et al., 2009). Et senere norsk studie på HUNT-materiale (Helseundersøkelsen i Nord-Trøndelag) av Ingrid Heuch, Heuch, Hagen, and Zwart (2016) på 18 068 nordmenn uten ryggsmarter viste derimot ikke evidens for en U-formet relasjon mellom fysisk aktivitet og risiko for ryggsmarter målt med spørreskjema. Det var heller ingen endring i risiko for ryggsmarter blant deltakerne med 3 timer hard fysisk aktivitet i løpet av en uke. Forfatterne konkluderer etter en oppsummering av lignende studier at risiko for ryggsmarter vil reduseres ved økende fysisk aktivitet på fritiden (Ingrid Heuch et al., 2016).

Oppsummert er ryggsmarter en ledende årsak til funksjonsbegrensninger, spesielt blant eldre. Ryggsmarter kan forstås i et biopsykososialt perspektiv hvor flere ulike faktorer spiller inn på graden av smerte og nedsatt funksjon. Det er usikkerhet knyttet til om det er en U-formet relasjon mellom fysisk aktivitet og grad av ryggsmarter, og hvorvidt ryggpasienter er mindre aktive enn den generelle befolkningen. Samtidig er det usikkert hvor stor effekt fear-avoidance kan ha på aktivitetsnivået til ryggpasienter.

2.2 Fysisk aktivitet

Fysisk aktivitet ble i 1985 av Caspersen et al. (1985) definert som enhver kroppslig bevegelse produsert av skjelettmuskler som resulterer i energiforbruk. Definisjonen er også benyttet av World Health Organization (WHO). Fysisk aktivitet i dagliglivet kan ytterligere kategoriseres etter arbeids- og sportsrelatert fysisk aktivitet (Caspersen et al., 1985). Selv om dette gjøres i litteraturen, blir fysisk aktivitet benyttet som et paraplybegrep på fysisk aktivitet på jobb, fritid og sportsrelatert aktivitet i denne oppgaven. Definisjonen som er redegjort for ovenfor betyr i praksis at den omfavner alle former for fysisk aktivitet.

Tidligere i oppgaven ble det blant annet trukket frem at fysisk aktivitet har en positiv effekt på kroniske smerter. Samtidig som man har en del studier på positive effekter av fysisk aktivitet er det og en del studier på hvilke effekter fysisk inaktivitet har på helsen. Grenseverdier for hva om betegnes som fysisk inaktiv baserer seg på anbefalingene på hvor fysisk aktiv man bør være

i løpet av en uke - for voksne er dette 150 minutter moderat fysisk aktivitet i uka, eller 75 minutter anstrengende aktivitet (WHO). Personer som har et aktivitetsnivå under disse anbefalingene kan derfor antas å være fysisk inaktive. I senere tid har mye av forskningen på fysisk aktivitet hatt hovedfokus på hvilke skadevirkninger fysisk inaktivitet har. Blant annet antas fysisk inaktivitet per 2008 å være ansvarlig for 9 % av prematur dødelighet globalt (I. M. Lee et al., 2012). Estimert på 9 % kan tenkes å øke fremover i takt med tendenser til en global trend hvor stadig flere er fysisk inaktive både på jobb og fritiden (S. W. Ng, Howard, Wang, Su, & Zhang, 2014; Shu Wen Ng & Popkin, 2012).

I en større internasjonal studie fra 2004 i ble aktivitetsnivået til over 50 000 deltakere fra 20 ulike land kategorisert som lavt, moderat eller høyt. Dataene var basert på et spørreskjema, der resultatene viste at prevalensen av høyt fysisk aktive varierte fra 21-63 %, og at aktivitetsnivået avtok med økende alder (Bauman et al., 2009). Nyere tall viser den samme tendensen, der aktivitetsnivået til deltakere fra 28 ulike EU-land ble undersøkt (Gerovasili, Agaku, Vardavas, & Filippidis, 2015). Lignende resultater ble også funnet i BioBank-studien, hvor akselerometermålinger av 96 600 deltakere fra Storbritannia viste en gradvis nedgang i aktivitet med økende alder (A. Doherty et al., 2017). Dette er i tråd med hva man har funnet i den siste undersøkelsen på fysisk aktivitet som er gjennomført i Norge (Helsedirektoratet, 2015). I undersøkelsen av aktivitetsnivået i Norge ble det benyttet akselerometer i 7 døgn som utgangspunkt for å beregne aktivitetsnivået til deltakerne. Resultatene fra undersøkelsen i Norge viste at aktivitetsnivået synker med økt alder, med et betydelig fall etter fylte 75 år (Helsedirektoratet, 2015). Evidensen som er presentert viser dermed at økt alder ledsages av mindre aktivitet på gruppenivå.

2.2.1 Måling av fysisk aktivitet

For enkelhetsskyld brukes begrepet metode i denne oppgaven som betegnelse på ulike måter å samle inn data på fysisk aktivitet – eksempelvis objektive metoder som akselerometer, eller subjektive metoder som spørreskjema. Måling av fysisk aktivitet uavhengig av metode har historisk sett vist seg å være utfordrende (Sylvia, Bernstein, Hubbard, Keating, & Anderson, 2014). Årsakene til hvorfor det er vanskelig å måle fysisk aktivitet er flere og utfordringene som er knyttet til aktivitetsmåling er avhengig av hvilken metode som benyttes. En felles utfordring for de ulike metodene er at de skal måle fysisk aktivitet som er regnet som en multidimensjonell og kompleks atferd (Falck et al., 2016). For å finne den meste egnede metoden for måling av aktivitet er det foreslått fem hovedelementer som bør vurderes i forkant

av måling: kvalitetene til aktiviteten (aktivitetstype, -intensitet, -hyppighet og -varighet), objektiviteten til dataene, deltakerbyrde (tid og innsats), kostnader og populasjonen (alder, kjønn, kroppsvekt, diagnose og komorbiditeter) (Sylvia et al., 2014).

Måling av fysisk aktivitet kan som nevnt gjøres gjennom flere ulike metoder. Definisjonen på fysisk aktivitet tilsier at man har to aspekter ved fysisk aktivitet, både bevegelse og grad av energiforbruk. Dette gir mulighet for å kunne måle fysisk aktivitet enten ved vurdering av fysisk aktivitet gjennom energiforbruk eller som atferd (Bahr & Helsedirektoratet, 2009, p. 117). Avhengig av om man ønsker å estimere energiforbruk eller fysisk aktivitet som atferd er det ulike måter å uttrykke disse to aspektene på. Litteraturen på feltet benytter begreper som Metabolic Equivalent Task (MET-verdier) og Energy Expenditure (EE) for måling av energiforbruk ved fysisk aktivitet. For fysisk aktivitet som atferd blir bevegelse målt og uttrykt som enten som rå akselerasjonsdata eller step-count som videre kan uttrykkes eksempelvis antall minutter lett, moderat eller anstrengende fysisk aktivitet (Sylvia et al., 2014). Videre utdypning rundt akselerasjonsdata følger nedenfor.

Metodene kan videre kategoriseres som enten subjektive eller objektive, ut fra hvordan data samles inn. Eksempelvis finnes subjektive metoder som selvrapporteringskjemaer og observasjon, eller objektive målinger som akselerometer, pedometer og hjertefrekvens-måling (Skender et al., 2016). Både subjektive og objektive metoder har svakheter som ytterligere gir utfordring i måling av fysisk aktivitet som utdypes i teorien nedenfor. Gullstandarden for måling av fysisk aktivitet er Doubly labeled water og omtales mer inngående i punkt 2.2.4.

Som nevnt er det knyttet en del utfordringer til metodene på måling av fysisk aktivitet. En konsekvens av den store variasjonen i metoder kompliserer reproduserbarheten og sammenfattingen av studier på fysisk aktivitet (P. H. Lee, Macfarlane, Lam, & Stewart, 2011; Skender et al., 2016). Dette fører til en del variasjon og diskusjon knyttet til standardisering av måling av fysisk aktivitet. En videre utdypning av utfordringer på subjektiv og objektiv måling av fysisk aktivitet følger nedenfor i teori-kapittelet og vil bli diskutert videre i diskusjons-delen av oppgaven.

2.2.2 Subjektiv eller selvrapportert måling av fysisk aktivitet

Det finnes mange ulike subjektive metoder: intervjuer, dagbøker og ulike typer spørreskjema (Westerterp, 2009). Blant de subjektive metodene er selvrapporteringsskjema den mest anvendte metoden (Sylvia et al., 2014). Fellesnevneren for selvrapporteringsskjema er at de er avhengige av respondentens evne til tilbakekalling av fysisk aktivitet – varierende ut ifra hvor langt tilbake skjemaet spør om. De ulike spørreskjemaene varierer i hva de måler, hvordan aktiviteten eller energiforbruket uttrykkes, kvaliteten til dataene og hvordan data er samlet inn (Sylvia et al., 2014). Eksempelvis ved bruk av IPAQ-SF måler man ikke en bestemt type aktivitet, men respondentene angir hvor mange minutter eller timer de gjennomsnittlig har tilbragt i våken tilstand i intensitetene sittende, gange, moderat eller meget anstrengende de siste 7 dagene.

Bruk av selvrapportert måling med skjema har sine klare fordeler ved at de er billige og enkle å administrere (Sylvia et al., 2014). Det faktum at de er billige og enkle i bruk kan tenkes å være en faktor på hvorfor det er en utstrakt bruk av selvrapporteringsskjema i befolkningsundersøkelser og intervensjonsstudier på eldre over 65 år (Falck et al., 2016). En annen fordel med selvrapporteringsskjema er at de i noen tilfeller kan gi detaljer omkring spesifikke typer aktivitet, eksempelvis slik IPAQ-SF gir informasjon om antall minutter per uke med gange.

Den klare ulempen knyttet til selvrapportering er bias ved tilbakekalling av egen fysisk aktivitet, også omtalt som recall bias (Sylvia et al., 2014). Evidens tyder på at faktorer som ryggsmerte og utdanning kan ha en påvirkning på hvor stor grad av bias det er knyttet til selvrapportering av fysisk aktivitet ved spørreskjemaesvarelsen. Studien av van Weering, Vollenbroek-Hutten, and Hermens (2011) viste at korrelasjon mellom selvrapportert fysisk aktivitet og akselerometermålt aktivitet korrelerte med -0.26 hos en gruppe på 32 ryggpasienter, sammenlignet med 20 friske voksne hvor korrelasjonskoeffisienten viste en verdi på 0.66 . En annen studie gjennomført i Danmark på 147 kontoransatte viste at blant de som hadde høy grad av korsryggsmerte også hadde økt grad av overrapportering av moderat til anstrengende aktivitet (MVPA) (Gupta, Heiden, Mathiassen, & Holtermann, 2018). Tegn på bias knyttet til utdanningsnivå fremkommer av en studie på International Physical Activity Questionnaire Long form (IPAQ-LF) som viste en signifikant forskjell i kriterievaliditet mellom høyt og lavt utdannede, hvor rapportering fra gruppen med lavt utdannede deltakere førte til en lavere

kriterievaliditet (Winckers et al., 2015). Det betyr i praksis at for unngåelse av skjevheter i datamateriale bør man inkludere deltakere med både høyt og lavt utdanningsnivå.

2.2.3 IPAQ-SF

IPAQ-SF er et spørreskjema som måler selvrapportert fysisk aktivitet de siste 7 dagene. Den korte versjonen, som blir benyttet i BACE-N, undersøker aktivitet som har vart 10 minutter eller mer innenfor de ulike intensitetene: gange, moderat- og meget anstrengende fysisk aktivitet (Craig et al., 2003). I 1996 startet et internasjonalt arbeid med å utvikle metoder for måling av fysisk aktivitet. Arbeidet ledet frem til en internasjonal konsensus-gruppe som utviklet International Physical Activity Questionnaire (Craig et al., 2003). Pilot-testing ble gjennomført året etter og man kom frem til to typer skjemaer i form av long form (31 items; IPAQ-LF) og short form (9 items IPAQ-SF). Disse to spørreskjemaene kan administreres enten via telefon eller selvrapportering. IPAQ er et av de mest brukte spørreskjemaene om fysisk aktivitet og er oversatt til mange ulike språk, deriblant Norsk (van Poppel, Chinapaw, Mookink, van Mechelen, & Terwee, 2010). Etter undertegnedes kjennskap er det kun utført en oversettelse, men ikke kryss-kulturell adaptasjon av den norske versjonen av IPAQ-SF.

For å minimere variasjon i hvordan man beregner fysisk aktivitet med IPAQ-SF er det utviklet en scorings-protokoll. Denne tar utgangspunkt i to måter for beregning av fysisk aktivitet, hvor fysisk aktivitet kan uttrykkes enten som kategorisk eller kontinuerlig variabel (www.ipaq.ki.se). Bruk av kategorisk variabel på IPAQ-SF gjøres ved summering av svarene til respondenten og deretter kategoriseres deltakerne som enten lav, moderat eller høy fysisk aktiv. Den kontinuerlige variabelen kan uttrykkes som Metabolic Equivalent task-minutter (MET-minutter) ved utregning av antall dager x antall minutter x MET-verdi (gange, moderat eller meget anstrengende). MET-verdier er et mål på estimering og klassifisering av energibruk ved ulike typer aktiviteter og eksempelvis for intensitetsnivåene til IPAQ-SF er MET-verdiene beregnet til henholdsvis 3.3, 4 og 8 MET for gange, moderat eller meget anstrengende fysisk aktivitet. 1 MET er for eksempel hvile, mens sykling kan regnes som 8.0 MET. Utregningen av MET følger formelen $3.5 \text{ mL} \times \text{kg}^{-1} \times \text{min}^{-1}$ (Ainsworth et al., 2011; Ainsworth et al., 2000).

For å undersøke psykometriske måleegenskaper til IPAQ-LF og -SF ble det i 2000 gjennomført validering og reliabilitetstesting i 14 ulike sentre fordelt på totalt 12 ulike land (Craig et al., 2003). Blant annet ble beregning av kriterievaliditet gjennomført på totalt 744 voksne hvor man

undersøkte korrelasjon mellom CSA akselerometer og begge versjoner av IPAQ (Short og Long form). Spearman korrelasjonskoeffisient mellom akselerometer og IPAQ viste en validitetsscore på henholdsvis 0.33 for den lange versjonen (IPAQ-LF), og 0.30 for den korte versjonen (IPAQ-SF) (Craig et al., 2003). Det ble i tillegg gjort reliabilitets-testing hvor man fant en reliabilitet på $\rho=0.76$ til IPAQ-SF. Like resultater på reliabilitet ble rapportert i en oversiktsartikkel av Silsbury, Goldsmith, and Rushton (2015) som inkluderte studiene av Craig et al. (2003) og Levy and Readdy (2009), hvor sistnevnte også benytter Spearman rho. Her er det viktig å bemerke at det ikke er anbefalt å bruke Spearman rho for undersøkelse av reliabilitet (Terwee et al., 2007).

I etterkant av estimering av måleegenskapene som ble gjennomført av Craig et al. (2003) har det blitt gjort en systematisk gjennomgang av enkeltstudier på validiteten til IPAQ-SF publisert 2011 av P. H. Lee et al. (2011) som viste en stor variasjon både i de inkluderte studiers metoder og resultater. Metodisk variasjon i de inkluderte studiene viste seg i form av ulike typer akselerometer, ulike prosesseringsmetoder og ulike grenseverdier i estimering av MET-verdier og energiforbruk (P. H. Lee et al., 2011). Resultatene fra den systematiske gjennomgangen viste en variasjon i kriterievaliditet fra 0.09 til 0.39 – noe som ifølge litteraturen ikke er godkjent nivå av validitet sammenlignet med minste akseptable korrelasjonskoeffisient på 0.70 ved sammenligning mellom to metoder (Terwee et al., 2007). Det påpekes i den systematiske gjennomgangen at det er behov for en standardisering av metodiske valg knyttet til kriteriestandarden ved validering av instrumenter for måling av fysisk aktivitet (P. H. Lee et al., 2011).

IPAQ er i senere tid modifisert til blant annet estimering av fysisk aktivitet blant eldre (www.ipaq.ki.se). International Physical Activity Questionnaire Elderly (IPAQ-E) ble validert opp mot akselerometer av Hurtig-Wennlof, Hagstromer, and Olsson (2010) hvor man fant korrelasjonskoeffisient varierende fra 0.277 – 0.471 målt ved Spearman rho. Negativ korrelasjon mellom IPAQ-E og akselerometer ble funnet både ved sammenligning av selvrappert sedatid og antall minutter tilbrakt i moderat fysisk aktivitet målt med akselerometer (Hurtig-Wennlof et al., 2010). En annen studie som sammenlignet totalt fysisk aktivitet målt med IPAQ-SF og akselerometer på henholdsvis 164 eldre menn og kvinner i Japan, fant Spearman rho hos menn på 0.42 og 0.49 for kvinner. For beregning av test-retest reliabilitet viste ICC-analyse 0.65 og 0.57 for menn og kvinner (Tomioka, Iwamoto, Saeki, & Okamoto, 2011). Til slutt viste valideringsstudien av IPAQ-SF av Grimm, Swartz, Hart, Miller,

and Strath (2012) hvor IPAQ-SF ble validert på 127 eldre med gjennomsnittsalder på 63.9 en korrelasjon på henholdsvis 0.042, 0.276 og 0.166 ved måling av moderat, gange og meget anstrengende fysisk aktivitet.

Oppsummert viser studier at de psykometriske egenskapene til IPAQ-SF er varierende med lav til akseptabel reliabilitet og gjennomgående lav grad av validitet. En av årsakene til den store variasjonen i resultater på validitet kan være knyttet til bruk av ulike kriteriestandarder benyttet i studiene (Garriguet, Tremblay, & Colley, 2015; Silsbury et al., 2015).

2.2.4 Objektiv måling av fysisk aktivitet

Objektive målemetoder inkluderer metoder som: akselerometer, pedometer, aktivitetsarmbånd og hjertefrekvensmåler (Sylvia et al., 2014). Per dags dato er *doubly labeled water* regnet som gullstandarden innen måling av fysisk aktivitet. Metoden estimerer energiforbruk ved fysisk aktivitet ved at to isotoper inntas oralt, for så å skilles ut av kroppen som enten vann eller CO₂. Ut fra mengden vann og CO₂ vil man da kunne estimere energiforbruket ved fysisk aktivitet (Westerterp, 2017). Denne måten å måle fysisk aktivitet på har imidlertid betydelige kostnader og har derfor blitt lite anvendt blant annet i metodestudier på validering av IPAQ-SF (P. H. Lee et al., 2011).

Intuitivt kan de store kostnadene knyttet til *doubly labeled water* tenkes å være en av årsakene til den nå vidstrakte bruken av akselerometer. Ved begynnelsen av 1980-tallet ble akselerometer ansett som en meget kostbart og komplisert metode. Etterhvert som at prisen, brukervennligheten og nøyaktigheten til akselerometre har bedret seg, har bruken av akselerometer siden 1990-tallet hatt en nærmest eksponentiell vekst innen forskning på fysisk aktivitet (Bassett, Troiano, McClain, & Wolff, 2015; R. P. Troiano, McClain, Brychta, & Chen, 2014).

Den mest åpenbare fordelen med bruk av akselerometer er at man utelukker bias tilknyttet tilbakekallelse av egen fysisk aktivitet (Clarke et al., 2017). Teknologiske fremskritt har tillegg ført til at akselerometrene nå er blitt mindre og lettere i bruk, noe som gjør at de blir mer anvendelig i større studier (Sylvia et al., 2014). En annen fordel med akselerometer er at dersom man er konsistent i hvilke typer akselerometer og prosesseringsmetoder som brukes vil det kunne gi reproducerbare data. En prospektiv kohortstudie målte fysisk aktivitet ved to

anledninger blant 203 kvinner i løpet av en 2-3 års-periode for å undersøke hvor reproduserbare akselerometer-data er. Studien brukte ActiGraph GT3X+ hvor deltakerne gikk med akselerometer i totalt to 7 dagers-intervaller. Forfatterne av studien fant en Intra Class Correlation fra 0.67 til 0.82 og konkluderte med at akselerometer gir reproduserbare data på både fysisk aktivitet og sedat tid (Keadle et al., 2017).

Parallelt med økt bruk av akselerometer i forskning på fysisk aktivitet, har det kommet nye typer akselerometer på markedet med ulike metoder knyttet til prosessering av akselerometer-dataene (Migueles et al., 2017). Derfor er det imidlertid en del uenighet innenfor fagfeltet om hvilket akselerometer som gir mest presise data, hvor på kroppen man skal plassere de, hvor mange akselerometer som trengs og ikke minst hvordan man skal prosessere rådataene (I. Cleland et al., 2013; R. P. Troiano et al., 2014). I senere tid har det vært en utvikling i hvordan man prosesserer akselerometer-data fra enklere algoritmer basert på grenseverdier over til mer kompliserte maskinlærings-algoritmer. Nyere evidens på disse prosesseringsmetodene viser at estimeringen av fysisk aktivitet varierer med 50 % mellom de ulike prosesseringsmetodene (Kerr et al., 2017). I tillegg til ulike variabler knyttet til prosessering av akselerometerdata er det også stor variasjon i hvordan parametere knyttet til akselerometre blir rapportert i ulike akselerometer-studier (Montoye, Moore, Bowles, Korycinski, & Pfeiffer, 2018). Den store mengden med data på både innsamling og prosessering av akselerometer-data gjør at man har ikke funnet en konsensus på hvilken metode som er best å benytte seg av (Migueles et al., 2017). En studie gjennomført på 80 friske menn viste at man ved å øke antall akselerometer fra en til to oppnår en statistisk signifikant bedring i nøyaktighet i måling av fysisk aktivitet. Studien viste imidlertid ingen videre forbedring i signifikans ved bruk av flere enn to akselerometer (I. Cleland et al., 2013). Et annet aspekt ved bruk av akselerometer er at de har en tendens til å underestimere energiforbruk når man eksempelvis går opp en trapp eller bærer tunge ting, særskilt ved måling på hoftområdet (Kerr et al., 2017; Lyden, Kozey, Staudenmeyer, & Freedson, 2011; Trost & O'Neil, 2014).

En ny utvikling i prosesseringen av akselerometerdata er bruken av maskinlæring for predikering og klassifisering av ulike typer aktivitet og intensitet. I en tidlig studie på maskinlæring av Ellis et al. (2014) ble 40 voksne, utstyrt med tre ActiGraph GT3X+ på begge hoftene og håndleddet. Deretter utførte deltakerne protokoll-styrte rutiner med ulike typer aktiviteter. For å predikere aktiviteten som ble gjennomført benyttet de seg av en random forest tree-algoritme. Dette er samme algoritme som er benyttet i studien av Willetts, Hollowell,

Aslett, Holmes, and Doherty (2018) og som videre blir benyttet i denne oppgaven. Resultatene fra studien til Ellis et al. (2014) var at predikering av fysisk aktivitet og intensitet med random forest tree-algoritme var mest nøyaktig med akselerometeret festet til håndleddet, med en total nøyaktighet på 80.2%.

Akselerometer har etterhvert blitt brukt i en rekke ulike større populasjonsstudier som mål på fysisk aktivitet. Et eksempel er BioBank-studien i Storbritannia hvor i overkant av 90 000 deltakere gikk med AX3 akselerometer over en periode på syv dager (A. Doherty et al., 2017). I Norge finnes det lignende folkehelseundersøkelser, eksempelvis Helseundersøkelsen i Nord-Trøndelag (HUNT4), der det ble tatt i bruk akselerometer som måleinstrument på fysisk aktivitet (<https://www.ntnu.no/web/hunt/hunt4>).

2.2.5 AX3 Akselerometer

Axivity AX3, som benyttes i denne oppgaven, er et triaksialt akselerometer på 23 x 32.5 x 7.6 mm og som veier 11 gram (www.axivity.com). Akselerometeret har innebygd 512 megabyte minne samt temperatur- og lys-sensor. Et oppladbart litium polymer-batteri driver akselerometeret. Akselerometeret er vann og støvbestandig tilsvarende IP6x/IP8x-standard og kan måle opptil 100 Hz med en sensitivitet mellom 2-16 G. For ordens skyld omtales akselerometeret som AX3 i denne oppgaven.

Det er gjennomført en del studier hvor man har sett på måleegenskapene til AX3. Nedenfor følger beskrivelser av noen studier hvor AX3 er benyttet som akselerometer. I en studie av Stewart et al. (2018) ble henholdsvis 42 barn og 33 voksne utstyrt med to AX3 akselerometer på lår og korsryggen for å måle nøyaktigheten til akselerometrene gjennom maskinlæring. Deltakerne i studien gjennomførte 10 ulike aktiviteter som eksempelvis sittende, liggende, stående, gående og løpende mens deltakerne ble observert i et laboratorie. Rådata fra akselerometrene ble så brukt til å utforme en algoritme for å gjenkjenne de ulike aktivitetene som tidligere ble gjennomført av deltakerne i laboratoriet. Algoritmens evne til å gjenkjenne de ulike aktivitetene, ble testet gjennom kryss-validering. Dette ble gjort både på målinger hvor de kombinerte data fra begge akselerometrene på lår og korsrygg og målinger fra hvert enkelt akselerometer. Resultatene fra studien viste at algoritmen hadde ved kombinasjon av begge akselerometrene en presisjon på 99.1% i gjenkjenning av 6 ulike aktiviteter (Stewart et al., 2018).

En studie hvor man brukte observasjon for å sammenligne tre ulike akselerometer AX3, APDM og ACTIGRAPH ved sakte gange på tredemølle viste at AX3 var mer nøyaktig i estimering av antall steg sammenlignet med de to andre akselerometrene. Studien inkluderte 25 personer i aldersspekteret 18-53, med gjennomsnittlig alder på 25 år (Feng, Wong, Janeja, Kuber, & Mentis, 2017). En annen valideringsstudie på AX3 med eldre mennesker med ulik varierende grad av funksjonsnedsettelse, gjennomførte validering ved å feste både AX3 og RT3 på håndledd, ankel, hofta og låret hvor korrelasjonsanalyser på rådata viste moderat korrelasjon (0.69) mellom AX3 og RT3 ved måling på håndleddet (Clarke et al., 2017).

BioBank-studien som har blitt gjennomført i Storbritannia bruker AX3 festet på håndleddet (A. Doherty et al., 2017). I den anledningen meldte de seg et behov for å vurdere nøyaktigheten til AX3 montert på håndleddet. Ved hjelp av en laboratorie-studie med aktivitets-protokoll og maskinlæring fant man at nøyaktigheten til håndledds-monterte AX3 akselerometer har akseptabel nøyaktighet i klassifisering av ulike former for aktivitet, sammenlignet med direkte observasjon. (K. Cooper et al., 2017).

Et annet studie som omhandlet feasibility til aktivitetsmåling med akselerometer hos eldre med proksimal femurfraktur, ble resultater fra tre ulike akselerometre (Fitbit, Misfit og Axivity AX3) festet på henholdsvis håndleddet, ankel og fremre lår sammenlignet (Schmal et al., 2018). Resultatene viste at AX3 er et egnet måleinstrument for denne pasientgruppen både med tanke på identifisering av wear-time og kategorisering av ulike intensitetsnivåer av fysisk aktivitet (Schmal et al., 2018).

For å måle anvendeligheten (feasibility) til AX3 ble det gjennomført en studie på 903 skolebarn hvor AX3 akselerometer ble festet til enten korsryggen eller til låret med tape. Beregning av wear-time ble fastslått med OMGUI. Resultatene fra undersøkelsen viste at gjennomsnittlig wear time beregnet over 10 dager med akselerometer var henholdsvis 7.29 og 6.80 dager for lår og korsrygg. Prosentandelen som fullførte over 9 dager med akselerometeret festet til kroppen viste 51.5% og 45.9% for henholdsvis lår og korsrygg (Schneller et al., 2017).

I punkt 2.2.4 ble det belyst teori omkring variasjon i prosessering av akselerometer-data. Ulike akselerometre krever ulike programvarer og for AX3 benyttes programvaren OMGUI. OMGUI er en åpen-kilde programvare som er bygd spesifikt for akselerometrene AX3 som er benyttet

i dette prosjektet (<https://github.com/digitalinteraction/openmovement/wiki/AX3-GUI>). Åpenkilde betyr i all hovedsak at den som har patent på programvaren har gjort kildekoden til programvaren åpen for alle som ønsker å gjøre endringer i koden. OMGUI er en slik programvare og er gratis. Per dags dato er dette eneste programvareløsning for å kunne overføre rådata fra Axivity AX3 akselerometer. Integret i OMGUI er det algoritmer som har sitt datagrunnlag fra ulike studier med ulike aldersgrupper som er gjennomført på enten AX3 eller lignende tri-aksiale akselerometer. Blant annet studien av Esliger et al. (2011) som danner grunnlag for en av prosesseringsmetodene benyttet i denne oppgaven. På lik linje med OMGUI er også programvaren som er benyttet i studiet av Doherty et al. (2017), og som brukes som prosesseringsmetode i dette prosjektet, også en åpen kildekode-programvare. Disse prosesseringsmetodene blir forklart i detalj i punktene 3.4.3-3.4.5 i metode-kapittelet.

Som nevnt tidligere har det nå blitt en økende trend med bruk av maskinlærings-algoritmer som kan predikere intensitet og ulike aktiviteter med rimelig nøyaktighet. I en studie gjennomført på 132 voksne og senere på 96 220 deltakere fra BioBank-studien i Storbritannia ble maskinlærings algoritme applisert på tri-aksiale rådata målt med AX3 akselerometer festet på håndleddet (Willettts et al., 2018). I studien ble det benyttet et fotokamera som tok bilder hvert 20. sekund, for å kunne kontrollere de dataene som ble produsert av maskinlæring og akselerometerdataene. Den totale klassifiseringsscoren med bruk av kappa inter-rater agreement viste en enighet mellom maskinlæringsmodellen og det bryst-monterte kameraet på 0.81 med en gjennomsnittlig nøyaktighet på 87 % (Willettts et al., 2018). Denne maskinlæringsalgoritmen er benyttet som prosesseringsmetode i denne oppgaven og omtales nærmere i punkt 3.4.5 i metode-kapittelet.

Oppsummert for måleegenskapene til AX3 akselerometer er at man ikke har validert akselerometeret opp mot gullstandarden doubly labeled water, men at studiene nevnt ovenfor tyder på akseptable måleegenskaper i form av validitet. I tillegg er det antydning til at AX3 er en anvendelig metode for måling av fysisk aktivitet.

2.3 Måleegenskaper

2.3.1 Validitet

Validitet er av Mokkink et al. (2010, s. 743) definert som graden av hvordan et instrument virkelig måler det konstruert det er ment til å måle. I form av begrepet konstruert, kan dette eksempelvis være fysisk aktivitet. Begrepet validitet kan videre deles inn i tre ulike subkategorier: innholds-validitet, kriterievaliditet, og konstruktvaliditet (de Vet, Terwee, Mokkink, & Knol, 2011, s. 150). Denne oppgaven har fokus på kriterievaliditet og derfor blir dette beskrevet nærmere i dette kapittelet.

Kriterievaliditet defineres som graden av hvordan skåresummen til et måleinstrument er en adekvat refleksjon av en gullstandard (Mokkink et al., 2010). Hva denne definisjonen impliserer er kravet om en gullstandard på det konstruert man skal måle. I form av hypotesetesting av kriterievaliditet, vil det derfor antas at måleinstrumentet man undersøker er like godt som gullstandarden (de Vet et al., 2011, s. 160). Man ønsker derfor å kunne si om dette er et valid verktøy for bruk i klinisk sammenheng – eksempelvis om IPAQ-SF er et valid mål på ryggpasienters fysiske aktivitet i løpet av en uke.

Kriterievaliditet kan ytterligere deles inn i begrepene prediktiv og samtidig validitet. Når man skal måle samtidig validitet er det en forutsetning at skåresummen av måleinstrumentet og gullstandarden måles på samme tidspunkt (de Vet et al., 2011, s. 159). Det innebærer i praksis at man får en nøyaktig sammenligning av tall fra IPAQ-SF og AX3, som i dette masterprosjektet ansees som gullstandard. Hvordan dette konkret gjøres i dette prosjektet utdypes nærmere i metode-kapittelet.

Beregning av kriterievaliditet gjøres vanligvis ved å beregne relasjon mellom to variabler, oftest i form av korrelasjonsanalyser, hvor gullstandarden og måleapparatet som skal vurderes representerer de to ulike variablene. Korrelasjonsanalyser gir en korrelasjonskoeffisient fra enten en positiv relasjon på 1 til negativ relasjon på -1, som gir en antydning til hvor sterk relasjon det er mellom gullstandarden og måleinstrumentet. Hva som regnes som akseptabel validitets-score kan ikke gis ut ifra en universell score, men en foreslått inndeling av korrelasjonskoeffisienter kan se slik ut: 00-.25 ingen til minimal korrelasjon, .26-.49 lav, .50-.69 moderat, 0.70-0.89 høy og .90-1.00 meget høy (Carter & Lubinsky, 2016, p. 323)

2.3.2 Feasibility

Feasibility er en form for undersøkelse av praktiske forhold rundt en intervensjon eller måling og ofte er noe som vurderes i forkant av RCT-studier, der man forsøker å si noe om hvorvidt den planlagte intervensjonen og de valgte målemetodene er gjennomførbare (Bowen et al., 2009). Ved måling av fysisk aktivitet har man behov for å vurdere om både deltakerne og forskerne i studien finner målemetodene som er benyttet som praktisk anvendbare.

Bowen et al. (2009) foreslår 8 punkter som en feasibility-studie bør fokusere på. Det første punktet som trekkes frem er *aksept* (acceptability). Altså en forutsetning om at dersom en intervensjon eller målemetode skal kunne brukes i praksis, må den kunne være akseptabel for både pasient/deltaker og måler/forsker. I studier på deltakere med helseproblemer er det essensielt at man ikke påfører deltakerne mer stress og ubehag enn nødvendig (Fitzpatrick, Davey, Buxton, & Jones, 1998). Utfallsmål på aksept kan for eksempel være hvor fornøyde deltakerne var med bruken, og om de kunne tenke seg å gjennomføre studien på ny (Bowen et al., 2009). *Etterspørsel* (demand) er og et aspekt som Bowen et al. (2009) tar opp som et mål på i hvilken grad deltakerne faktisk bruker intervensjonen som er planlagt. Et potensielt utfallsmål på etterspørselen er eksempelvis i hvilken grad deltakerne har brukt spørreskjemaet eller aktivitetsarmbåndet. *Implementering* (implementation) sier noe om i hvor stor grad intervensjonen/målingen kan implementeres i en ikke fullt kontrollert kontekst. I dette legger de hvorvidt intervensjonen/målingen kan legges til en klinisk setting eller eksempelvis i en kontekst hvor man ikke har like stor kontroll som i en vitenskapelig studie eller et laboratorie. Mulige utfallsmål som nevnes er i hvor stor grad man klarer å fullføre de målingene man har planlagt i forkant av forsøket. *Praktiske* (practicality) årsaker nevnes som i hvilken grad man med eksisterende ressurser og omkringliggende forhold kan gjennomføre intervensjon/målingen utenfor selve studiet. Her nevnes eksempelvis en kost-analyse som et mulig utfallsmål. *Adaptasjon* (adaptation) kan brukes som et begrep på hvor anvendbar intervensjonen/målingen er i et annet format eller på en annen populasjon, hvor potensielle utfallsmål er om man oppnår lignende resultater i andre lignende studier. *Integrering* (integration) innebærer om den nye intervensjonen/målingen kan implementeres inn i et allerede eksisterende system hvor man må vurdere kostnadene. *Ekspansjon* (expansion) omhandler i hvor stor grad det man har undersøkt kan utvides og danne nye programmer, eller tilbud. Her kan aktuelle utfallsmål være i hvilken grad intervensjonen/måling passer inn rent økonomisk eller kulturelt. Til slutt er *begrenset effektivitet* (limited efficiency) et begrep som

omhandler i hvor stor grad den nye intervensjonen/målingen viser tendens til å kunne være suksessfull i den populasjonen man har valgt. Her er et mulig utfallsmål å se på selve effekten av intervensjonen/måling (Bowen et al., 2009).

Som beskrevet ovenfor er det mange mulige faktorer og overveielser man potensielt kan ta i betraktning i et feasibility-studie. Begrepene beskrevet ovenfor er ingen fasit på hva et feasibility-studie inneholder, men kan gi noen føringer på hvilke faktorer som er viktig å vurdere.

3.0 Metode

3.1 Design

Dette masterprosjektet er en metodestudie for å undersøke den samtidige validiteten til spørreskjemaet International Physical Activity Questionnaire Short Form (IPAQ-SF) med AX3 akselerometer på ryggpasienter over 55 år. Som gullstandard for validering av IPAQ-SF benyttes Axivity AX3 akselerometer. Det er i tillegg gjennomført en feasibility-studie for å undersøke deltakernes erfaring med IPAQ-SF og akselerometeret, samt gjennomførbarheten til disse metodene. Masterprosjektet, både ved validerings- og feasibility-studien, har et tverrsnitts-design.

Utgangspunktet for denne studien var å undersøke både validitet og reliabilitet til IPAQ. Grunnet lav responsrate (n=10), ble det avgjort å utelukke reliabilitets-testing. Det var videre ønskelig med undersøkelser rundt potensielle over-/underrapporteringer i IPAQ på tvers av kjønn og utdanningsnivå. Dette ble heller ikke gjennomført grunnet det lave utvalget i dette prosjektet.

3.2 Rekruttering og utvalg

Deltakere fra BACE-N-studien ble rekruttert til dette masterprosjektet. Årsaken til dette var på grunnlag av ønske fra prosjektledelsen om å undersøke samtidig validitet til IPAQ-SF med AX3 akselerometer. Fordi utvalget i dette masterprosjektet er rekruttert fra deltakere BACE-N-studien er inklusjons- og eksklusjonskriteriene gjeldende også for dette prosjektet. Inklusjonskriteriene for BACE-N er som følger: kvinner og menn som er 55 år og eldre som oppsøker primærhelsetjenesten (fastlege, fysioterapeut eller kiropraktor) med en ny episode med rygg smerter (innledet av 6 måneder uten konsultasjon for lignende plager) vil rekrutteres. Eksklusjonskriteriene er: pasienter med kognitive nedsettelse som hindrer dem i å fullføre studiens spørreskjemaer eller dersom de har vanskeligheter med å snakke og skrive norsk, eller dersom de har betraktelig nedsatt syn. Pasienter som har betraktelige mobilitets-vansker (for eksempel benyttelse av rullestol) vil bli ekskludert dersom de har vansker med å fullføre fysisk undersøkelse. BACE-N rekrutterer pasienter fra flere geografiske områder i Norge. Fordi det i dette masterprosjektet var lagt opp til et personlig møte mellom deltakere og masterstudenten

ble det for masterprosjektet lagt til et ytterligere eksklusjonskriteria som sa at pasientene måtte befinne seg på Østlandsområdet.

Utvalget i dette prosjektet ble rekruttert fra databasen til BACE-N ved at det ble sendt ut invitasjon til deltakelse i masterprosjektet i forbindelse med 3, 6, 12 eller 24 måneds oppfølging i BACE-N. Oppfølging i BACE-N ved disse tidspunktene består i elektronisk utfylling av spørreskjemaer. Deltakerne fikk skriftlig påminnelse enten via SMS eller e-post om utfylling av spørreskjemaer i BACE-N, samt en elektronisk lenke til nevnte spørreskjemaer. Samtidig med innkalling til oppfølging og elektronisk utfylling av spørreskjemaer i BACE-N ble det sendt informasjon om dette masterprosjektet med en kort beskrivelse av mål og hensikt med studiet og invitasjon til å delta. I desember 2018 ble det i tillegg sendt ut en julehilsen til deltakere i BACE-N i Oslo-området i brevform som tiltak for å øke responsraten i masterprosjektet. Kontaktinformasjon til undertegnede ble lagt ved og pasientene fikk mulighet til å ta kontakt direkte med masterstudent eller rekrutteringsansvarlig for BACE-N. Majoriteten av deltakerne ble rekruttert i Oslo og omegn. Interesse fra deltakere ble meldt enten til stipendiat i BACE-N-prosjektet eller til masterstudenten direkte.

Etter at kontakt var etablert ble det avtalt et fysisk møte på et sted det var enkelt for BACE-N-deltakere å treffes. I møtet med masterstudenten og deltakeren ble det gitt utdypende muntlig og skriftlig informasjon om prosjektet, og informert samtykke for deltakelse i masterprosjektet ble undertegnet av deltaker og masterstudent (se også punkt 3.3). Under møtet med deltakerne ble IPAQ-SF-skjemaet gjennomgått og deltakerne ble nøye instruert i å svare så konsistent som mulig på IPAQ-SF-skjema og unngå svaralternativet «vet ikke» på IPAQ-SF-skjemaet (vedlegg 1). Dette ble gjort for å unngå tapte data fra IPAQ-SF. På slutten av møtet ble deltakerne utstyrt med akselerometer som skulle bæres på ikke-dominant hånndledd i 7 døgn. Videre beskrivelser av hvilke parametere og innstillinger som ble benyttet i innhenting av akselerometer-data følger under punkt 3.4.2. Fordi IPAQ-SF er et spørreskjema som måler selvrapportert fysisk aktivitet de siste 7 dagene, ble deltakerne bedt om å fylle ut IPAQ-SF-skjema etter å ha gått 7 døgn med akselerometer, for så å legge både skjemaet og akselerometer i en forhåndsfrankert konvolutt. Den ferdigfrankerte konvolutt var adressert til Statens Arbeidsmiljøinstitutt (STAMI) hvor masterstudenten hadde kontor under arbeidet med dette masterprosjektet.

3.3 Etikk

Som utgangspunkt for etisk forskning ble dette prosjektet gjennomført i henhold til Helsinkideklarasjonen ("[The Helsinki Declaration of the World Medical Association (WMA). Ethical principles of medical research involving human subjects]," 2014).

BACE-N prosjektet er fremlagt for Regional Etisk Komite (REK), men klassifisert som kvalitetssikring og dermed et prosjekt som faller utenom helseforskningslovens virkeområde (ref 2014/1634). BACE-N har godkjenning fra Norsk senter for forskningsdata (NSD) (referansenummer: 42149). For å kunne gjennomføre masterprosjektet slik det var skissert var det imidlertid nødvendig med endringsmelding tilknyttet BACE-N-prosjektet til NSD. Et samtykkeskjema med utgangspunkt i mal fra NSD ble utformet og godkjent av NSD i forkant av utsendelse til aktuelle deltakere i masterprosjektet (vedlegg 3).

Ved interesse for deltakelse ble et møtetidspunkt og -sted avtalt enten per telefon eller tekstmelding. Alle deltakerne fikk lese gjennom samtykkeinformasjon ved starten av møtet og informert samtykke ble underskrevet av pasient og undertegnede før datainnsamling tilknyttet masterprosjektet ble påbegynt. I tillegg ble det gitt utdypende muntlig informasjon fra student om masterprosjektets omfang og mål. Under møtet fikk deltakerne mulighet til å stille spørsmål til student omkring prosjektet. Den underskrevne samtykkeinformasjon ble oppbevart i et låst arkiv på STAMI. Alle data ble anonymisert med ID nummer slik at de ikke kunne gjenkjennes. Dataene ble lagret på passord-beskyttet privat datamaskin.

3.4 Variabler

3.4.1 Deskriptive variabler

For å beskrive utvalget i dette prosjektet ble det hentet et utvalg med aktuelle variabler fra datasamlingen til BACE-N. Dette var deskriptive baseline-variabler som alder, kjønn, arbeid og utdanningsnivå. I tillegg ble data for kartlegging av ryggsmertor og funksjonsnivå tilgjengeliggjort både ved baseline (inkludert til BACE-N) og ved oppfølgingstidspunktet i BACE-N (0, 3, 6, 12 eller 24 måneds oppfølging) hvor deltakerne ble rekruttert til dette masterprosjektet. Data på fear-avoidance og livskvalitet var kun tilgjengelig fra baseline (inkludert til BACE-N). Variablene ble målt spesifikt i form av ryggsmerteintensitet målt ved

Numeric Rating Scale, funksjon i dagliglivet målt ved den norske validerte versjonen av Roland Morris Disability Questionnaire, fear-avoidance målt med norske Fear-Avoidance Beliefs Questionnaire – fysisk aktivitetskomponent og livskvalitet målt med den norske versjonen av SF-36 (Downie et al., 1978; Grotle, Brox, & Vollestad, 2003, 2006; Roland & Morris, 1983; Waddell, Newton, Henderson, Somerville, & Main, 1993). Kartlegging av disse variablene er i tråd med anbefalinger på hvilke utfallsmål som bør bli benyttet i studier av ryggpasienter (Chiarotto et al., 2018).

3.4.2 International Physical Activity Questionnaire (IPAQ-SF)

Hovedfokus for dette masterprosjektet er validering av selvrapporterte aktivitetsdata fra IPAQ-SF, som er det samme skjemaet som benyttes for kartlegging av fysisk aktivitet i BACE-N. Siden dette masterprosjektet opprinnelig skulle undersøke både validitet og reliabilitet (se punkt 3.1) svarte deltakerne på to skjemaer hver (ett da de fikk utdelt akselerometer og ett ved innlevering 7 dager senere), men grunnet at det ikke ble gjennomført reliabilitets-testing ble kun ett skjema (utfylt ved innlevering av akselerometer) benyttet i dette prosjektet.

Dataene fra IPAQ-SF-skjemaene ble scoret i henhold til scorings-protokollen ad 2005 (www.ipaq.ki.se) med en modifikasjon på utregning av aktivitetsdata hvor MET-verdiene ble utelukket. All scoring av IPAQ-SF-skjemaene ble gjort manuelt ved bruk av kalkulator. Standard prosedyre for scoring av IPAQ-SF jamfør protokoll fra 2005 er at svarene fra IPAQ skulle scores med følgende utregning: «*x dager * x minutter * 3.3, 4, eller 8 (gående, moderat, meget anstrengende) MET*». Dette gir en sum i antall MET-minutter. Som nevnt tidligere i teorien er MET et mål på aktivitetsintensitet, og gir altså ikke en score i hvor mange minutter deltakeren har vært i enten gange, moderat eller meget anstrengende fysisk aktivitet. For sammenligning av data fra IPAQ-SF og akselerometer var det derfor nødvendig å endre på utregningen av IPAQ-SF-data, slik at utregningen ble slik: «antall dager * antall minutter». Dette ga en verdi på hvor mange minutter deltakerne oppga å være i en av intensitetene målt av IPAQ-SF og muliggjorde korrelasjonsanalyser med data fra akselerometer. Slik utregning av fysisk aktivitet på data fra IPAQ-SF har også blitt gjort av Sanda et al. (2017) i en valideringsstudie av IPAQ-SF på gravide kvinner.

Vasking og trunkering av data fra IPAQ-SF ble gjort i henhold til protokollen (www.ipaq.ki.se). Det innebærer at der deltakerne hadde svart mer enn 3 timer (180 minutter) gjennomsnittlig aktivitet på en av de tre ulike intensitetene gange, moderat eller meget anstrengende, ble disse

svarene trunkert til 180 minutter. Dette ble gjort siden man i scorings-protokollen tar utgangspunkt i at man kan ha maksimalt 21 timer med aktivitet ukentlig i hver av de respektive aktivitets-intensitetene. Der hvor noen hadde svart 1-2 timer ble det tatt gjennomsnitt av svaret for å kunne ta med disse verdiene videre i scoring av skjemaet. De skjemaene hvor deltakerne hadde svart «vet ikke» på enten gange, moderat eller meget anstrengende ble ekskludert fra videre analyser. Til slutt ble hver av disse scoresummene ført inn i eget Excel-ark med fordeling av antall minutter i hver av de tre ulike intensitetene på gange, moderat og meget anstrengende.

3.4.3 Axivity AX3 akselerometer

IPAQ-SF er i dette prosjektet validert med akselerometeret AX3 for å kunne undersøke den samtidige validiteten til IPAQ-SF. Rådata fra akselerometer ble prosessert ved bruk av to ulike metoder (Metode1_{Esliger} og Metode2_{Doherty}). Disse representerer henholdsvis en klassisk fremgangsmåte med grenseverdier for klassifisering av aktivitet innebygd i programvaren OMGUI (punkt 2.2.5) og en nyere metode hvor maskinlæring brukes til å predikere aktivitet. Beskrivelse av måleegenskapene til Axivity AX3 er angitt i teori (punkt 2.2.5).

For å lettere kunne skille mellom de to ulike prosessering-metodene benyttes begrepene Metode1_{Esliger} og Metode2_{Doherty} konsekvent gjennom oppgaven. Metode1_{Esliger} representerer prosesseringsmetoden bygget på grenseverdier fra Esliger et al. (2011) som inngår i den tidligere nevnte programvaren OMGUI (punkt 2.2.5), mens Metode2_{Doherty} bygger på en nylig utviklet maskinlærings-algoritme av Willetts et al. (2018) som blir benyttet på data fra BioBank-studien i Storbritannia (A. Doherty et al., 2017). De samme rådataene ble benyttet i både Metode1_{Esliger} og Metode2_{Doherty}.

Deltakerne ble utstyrt med AX3 akselerometer plassert i klokkereim av silikon på det ikke-dominante håndleddet som de skulle gå med i syv hele døgn. For å øke måletiden samt forebygge tapte måledata fra akselerometer, ble ulike tiltak iverksatt. Deltakerne ble oppmuntret til å bære akselerometrene kontinuerlig de 7 døgnene, med unntak av aktiviteter under vann. Akselerometrene ble forhåndsinnstilt til å måle over en periode på 7 døgn. Akselerometrene ble også innstilt slik at de skulle blinke underveis i målingen slik at deltakerne skulle kunne forstå dersom teknisk feil hadde oppstått og dermed ta kontakt med studenten for å bli utstyrt med nytt akselerometer.

Samtlige akselerometer ble innstilt på 100 Hz målefrekvens av mulige 12.5-3200 Hz. Dette ble gjort siden det er anbefalt i den systematiske oversiktsartikkelen av Migueles et al. (2017) og ifølge produsenten skal bruk av 100 Hz kunne gi en batterilevetid på 14 dager (www.axivity.com), noe som gir margin med tanke på måling av aktivitet over 7 døgn. Av de ulike alternativene på 2, 6, 8 og 16 i form av sensitivitet, ble det valgt ± 8 G sensitivitet. For å kunne videre fastslå hva som er regnet som en valid måle-dag og uke med akselerometer ble det tatt utgangspunkt i oversiktsartikkelen av Migueles et al. (2017) hvor anbefalt valid dag hos både voksne og eldre voksne regnes som ≥ 10 timer og valid uke ≥ 4 dager.

Etter at akselerometrene ble mottatt i retur fra deltakere ble disse koblet til en Windows-kompatibel datamaskin via USB. Siden man utelukkende kan laste ned rådata fra AX3 med programvaren OMGUI (punkt 2.2.5) ble denne programvaren benyttet til nedlastning av rådata fra akselerometrene. En av funksjonene i OMGUI er at man kan få en grafisk fremstilling av rådataene presentert i aktogram, som gir mulighet til å vurdere aktivitetsnivået på et overordnet og generelt nivå. Inspeksjon av hver enkelt CWA-fil med rådata gjorde at man fikk vurdert om eventuelle tekniske feil hadde avbrutt målingen, noe det ikke hadde. Videre ble hver enkelt CWA-fil med rådata lastet ned for å kunne gjøre videre analyser i form av applisering av støyfilter, wear-time validering og klassifisering av fysisk aktivitet.

Siden de to ulike metodene benyttet på prosessering av akselerometerdata er delt inn i ulike intensiteter ble det gjennomført korrelasjonsanalyse på de variablene som matchet med intensitetene fra IPAQ-SF (gange, moderat og meget anstrengende). Blant annet var det kun Metode2_{Doherty} som hadde gange som egen intensitet, og dermed den eneste til å kunne undersøke korrelasjon med IPAQ-SFs gange. På den andre enden av intensitetsskalaen var meget anstrengende en intensitet man kun fikk hentet ut fra Metode1_{Esliger}. Derfor var moderat intensitet den eneste av IPAQ-SF-intensitetene hvor man fikk gjennomført korrelasjonsanalyser med begge prosesseringsmetodene.

3.4.4 Metode1_{Esliger} Wear-time validering, kvantifisering og klassifisering av rådata til aktivitetsdata

I forkant av prosessering av rådata ble det applisert et støyfilter på 0.5-20 Hz for å redusere digital støy i målingene til akselerometrene. Støyfilteret på 0.5-20 Hz er standard i OMGUI, og er i tillegg tidligere benyttet i en valideringsstudie på AX3 akselerometer (Clarke et al., 2017). Deretter ble det gjennomført wear-time validering gjennom programvaren OMGUI, som er

standard wear-time validering i OMGUI. Algoritmen på wear-time validering bygger på en studie av energiforbruk med akselerometer festet på håndledd hos gravide og ikke-gravide kvinner (van Hees et al., 2011). Akselerometrene som er benyttet i studien til van Hees et al. (2011) tar utgangspunkt i GENE A akselerometer, som i likhet med AX3 er et tri-aksialt akselerometer. Valideringen på om deltakerne hadde på seg akselerometeret eller ikke ble gjort gjennom standard deviasjon og verdi-spekteret som ble undersøkt i bolker på 30 minutter. En bolk på 30 minutter som viser en standard deviasjon på mindre enn 3.0 milligravity ($1 \text{ mg} = 0.00981 \text{ m} \cdot \text{s}^{-2}$), eller dersom verdi er mindre enn 50 milligravity på minst to av tre akser regnes som at deltakeren ikke hadde på seg akselerometeret (van Hees et al., 2011). Teorien om at et akselerometer er tatt av når to av tre akser viser en standard deviasjon mindre enn 3.0 milligravity kommer fra van Hees et al. (2011) egne eksperimenter hvor de har funnet ut hvor mye digital støy som produseres fra akselerometrene når de ligger helt i ro. Eksperimentet viste at det digitale støyet fra akselerometeret produserer utslag i form av 2.6 mg standard deviasjon ved total stillstand. Derimot er kunnskapsgrunnlaget på at 50 mg er å regne som non-wear time fra en studie hvor bi-aksiale akselerometer, festet til brystkassen til deltakerne av studien, viste at respirasjonssyklusen alene har en bevegelse på 10 mg, mens bevegelsen til et hjerteslag gir et utslag på 80 mg (Phan, Bonnet, Guillemaud, Castelli, & Pham Thi, 2008). Sluttresultatet av wear-time validering produserte et Excel-ark som viste hele måleperioden fordelt på 1-timers bolker med verdiene 1 eller 2. Verdien 2 regnes som at deltakeren har hatt på seg akselerometeret kontinuerlig i løpet av denne timen, mens 1 regnes som at deltakeren har hatt på seg akselerometeret den første halvtimen og deretter tatt akselerometeret av.

Etter validering av wear/non wear time ble selve analyseringen av rådata fra akselerometer over til aktivitetsdata gjennomført. For å kunne prosessere disse dataene ble det tatt utgangspunkt i tidligere nevnte OMGUI. OMGUI har som nevnt ulike algoritmer basert på ulike studier. Siden man i denne studien blant annet hadde valgt å feste akselerometeret på håndleddet og hadde et eldre utvalg ble algoritmer basert på Esliger et al. (2011) benyttet. Esliger et al. (2011) gjennomførte en validering av GENE A akselerometer på 60 menn og kvinner i alderen 40-65 år med grenseverdier inndelt etter MET-verdier for kvantifisering av intensitetene sedat (>1.5 MET), lett (1.5-3.99 MET), moderat (4.00-6.99 MET) og anstrengende fysisk aktivitet (7+ MET). Disse grenseverdiene og intensitetsnivåene utgjør det som er Metode_{Esliger} i dette prosjektet.

Resultatet av gjennomført prosessering i OMGUI var en CSV-fil innholdende en uke med antall minutter i intensitetsnivåene sedat, lett, moderat og anstrengende i 1-minutts epoker. Hver enkelt av intensitetsnivåene ble summert til totalt antall minutter aktivitet for hver av de fire ulike intensitetsnivåene (sedat, lett, moderat og anstrengende fysisk aktivitet). Tallene fra de ulike intensitetsnivåene ble så ført inn i en samlet Excel-fil sammen med resultater fra IPAQ-SF-skjemaer og Metode2_{Doherty}.

3.4.5 Metode2_{Doherty} Wear time validering, kvantifisering og klassifisering av rådata til aktivitetsdata

En nyere metode for prosessering av rådata fra akselerometer, som blir omtalt som Metode2_{Doherty}, er hentet fra en metode utviklet av A. Doherty et al. (2017) og Willetts et al. (2018) for UK BioBank-studien og er i dette prosjektet som en alternativ måte å prosessere rådata fra akselerometrene. En maskinlæringsalgoritme ble utviklet Willetts et al. (2018) på akselerometerdata fra 90 000 deltakere fra BioBank-studien til A. Doherty et al. (2017). Dette er en nyere form for prosessering av akselerometer-data som er utviklet ved å benytte AX3 akselerometer montert på håndleddet og som kan gi aktivitetsdata på blant annet minutter med gange (Willetts et al., 2018). For å kunne benytte seg av programvaren ble en datamaskin med Linux benyttet som utgangspunkt for å prosessere rådataene fra akselerometrene. Dette var nødvendig på grunn av kompatibiliteten til programvaren som utelukkende kunne kjøre på Linux som operativ-system. Programvaren som ble brukt for å analysere dataene ble lastet ned fra <https://github.com/activityMonitoring/biobankAccelerometerAnalysis>. Programvaren er kommando-basert som kjøres via Python 3. Integret i programvaren er både filtrering av støy, wear-time validering og klassifisering og kvantifisering av akselerometerdata over til aktivitetsdata.

For å beskrive non-wear time blant deltakerne har programvaren tatt utgangspunkt i stasjonære perioder over minst 60 minutter hvor standard deviasjon av alle tre aksene viser en akselerasjon på mindre enn 13.0 milligravity (A. Doherty et al., 2017). Alle perioder med non-wear time ble loggført av programvaren som videre ble lagret i en egen CSV-fil. Hos de deltakerne som ikke hadde tatt av akselerometeret underveis i måleperioden var denne CSV-filen med non-wear time tom og helt uten talldata. Ytterligere ble det produsert en tekstfil i form av JSON-fil med gjennomsnittlig verdi på antall dager wear-time med akselerometeret. Den gjennomsnittlige verdien på antall dager med wear-time ble ekstrahert fra JSON-filen og flyttet til eget Excel-ark.

Hver av CWA-filene med rådata ble prosessert med programvaren og som resultat fikk man ut en komprimert CSV-fil med antall minutter i de ulike intensitetsnivåene søvn, sedat, lett, gange og moderat i en oppløsning på 30-sekunder epoker. Det ble gjort omregning av aktivitetsdataene til hele minutt slik at de var sammenlignbare med både IPAQ-SF og Metode1_{Esliger}.

Aktivitetsdata, i form av minutter aktivitet per uke, fra både Metode1_{Esliger} og Metode2_{Doherty} ble ført inn i eget Excel-ark for videre analyser i SPSS. Utgangspunktet for å kunne gjøre korrelasjonsanalyser på de to ulike prosesseringsmetodene var totalt 9 ulike intensitetsnivåer med aktivitet og sedat tid fordelt på de to ulike prosesseringsmetodene. For Metode1_{Esliger} var dette intensitetsnivåene sedat, lett, moderat, og anstrengende. Metode2_{Doherty} produserte data på totalt 5 ulike intensitetsnivå i form av sedat, søvn, lett, gange og moderat. For å skape et sammenligningsgrunnlag for de ulike intensitetene ble data fra lett fysisk aktivitet og gange i Metode2_{Doherty} summert sammen for å kunne matche lett fysisk aktivitet i Metode1_{Esliger}. Årsaken til dette ble gjort på grunnlag av grenseverdiene til Metode1_{Esliger} hvor lett-intensitetene hadde en grenseverdi på 1.5-3.99 MET, tilsvarende hadde Metode2_{Doherty} klassifisert lett fysisk aktivitet som 2.2 MET og gange som 3.2 MET. Det var derfor grunnlag for å slå sammen lett og gange i Metode2_{Doherty} for å styrke grunnlaget for sammenligning. For å kunne gjøre korrelasjonsanalyser på Metode1_{Esliger} og Metode2_{Doherty} ble data som ikke var behandlet annet enn summering av totalt antall minutter i de ulike intensiteter benyttet (lett, gange, moderat, meget anstrengende).

Til slutt ble intervaller på 10 minutter eller mer ble ekstrahert fra aktivitetsdataene fra begge prosesseringsmetodene for å kunne gi et sammenligningsgrunnlag med dataene fra IPAQ-SF (punkt 2.2.3). Dette ble gjort manuelt ved å hente ut data fra hvert enkelt sett med aktivitetsdata fra begge prosesseringsmetoder og deretter summere intervaller på >10 minutter for intensitetsnivåene gange, moderat og meget anstrengende fysisk aktivitet. Dette ble gjentatt to ganger for å minske faren for tapte data eller feilberegning.

3.5 Feasibility

For å kunne vurdere aspekter ved feasibility til både IPAQ-SF og AX3 ble det utformet et spørreskjema. Spørreskjemaet (vedlegg 2) inneholdt totalt ti spørsmål basert på spørsmål

benyttet i tre tidligere studier hvor feasibility ved bruk av akselerometer-måling på pasienter med neurologisk dysfunksjon, pasienter med bronkiektasier og tenåringer er undersøkt (Hale, Pal, & Becker, 2008; O'Neill et al., 2017; Scott et al., 2017). Spørsmålene ble tilpasset utvalget på 55+ (Hale et al., 2008; O'Neill et al., 2017; Scott et al., 2017). Spørreskjemaet var todelt hvor den ene halvdel omhandlet spørsmål knyttet til bruk av akselerometeret, mens den andre halvparten omhandlet spørsmål knyttet til bruk av IPAQ-SF-skjemaet, hvor av totalt 2 spørsmål knyttet til akselerometer var egenproduserte og spørsmålene knyttet til IPAQ-SF-skjemaet var adapterte fra akselerometer-spørsmålene (vedlegg 2). En 5-punkts Likert-skala ble benyttet som svaralternativ til spørsmålene. Spørreskjemaet ble utformet midtveis i prosjektet da man erfarte svært lav responsrate på masterprosjektet og dessuten så at det ville være hensiktsmessig med data på feasibility knyttet til IPAQ-SF og AX3. Av den grunn ble deltakere som allerede var inkludert i masterprosjektet (n=5) kontaktet med spørsmål om utfylling av spørreskjema for feasibility via e-post. Nyrekrutterte deltakere ble bedt om å returnere skjemaet sammen med akselerometeret og IPAQ-SF-skjemaet.

Data fra wear-time valideringen fra begge prosesseringsmetoder ble benyttet ved å sette opp median antall dager wear-time i løpet av 7 døgn med akselerometermåling. Hver av de to ulike prosesserings-metodene hadde ulike mål på wear-time og det ble derfor gjort omregninger slik at man uttrykt dette som median antall dager wear-time med akselerometer.

3.6 Statistikk

SPSS versjon 25 for Mac ble benyttet som programvare for all statistisk beregning som ble gjennomført i dette masterprosjektet.

3.6.1 Deskriptiv statistikk

Kategoriske variabler for beskrivelse av utvalget er presentert som andeler og prosent. Variabler ansett som kontinuerlige (RMDQ, NRS, SF-36 og FABQ) ble vurdert hvorvidt de var normalfordelt ved inspeksjon av dataene i et histogram med normalfordelingskurve. Siden utvalget var lite ble variablene behandlet som ikke-normalfordelte, ble derfor median samt interkvartilbredde benyttet i deskriptive analyser.

3.6.2 Samtidig validitet

For å fastslå hva som er akseptabel verdi for korrelasjon mellom to metoder er det foreslått ulike verdier. En inndeling som er foreslått og som vil bli benyttet i denne oppgaven ser slik ut: 00-.25 ingen til minimal korrelasjon, .26-.49 lav, .50-.69 moderat, 0.70-0.89 høy og .90-1.00 meget høy korrelasjon (Carter & Lubinsky, 2016, p. 323). For dette prosjektet ble det fastslått en cut-off verdi på 0.70 som akseptabel korrelasjon mellom IPAQ-SF og akselerometer (Terwee et al., 2007). Signifikansnivå ble satt til 0.05. Ved gjennomføring av korrelasjonsanalyser for validering av IPAQ-SF ble det, grunnet ikke-normalfordelte data og lite utvalg (n=10), benyttet Spearman rang-analyse.

3.6.3. Feasibility

Dataene fra spørreskjemaet på feasibility (vedlegg 2) ble plottet inn i et Excel-ark hvor median ble angitt for hvert enkelt spørsmål sammen med minimums- og maksimumsverdi. Dette ble gjort for å presentere variasjonen i hvordan deltakerne responderte på de ulike spørsmålene. Dataene som fremkom blir presentert med tabell i resultat-kapittelet.

Dataene på wear-time som fremkom fra de to ulike prosesserings-metodene ble satt opp i eget Excel-ark for utregning av median antall dager wear-time for begge prosesseringsmetodene. Grunnet mangel på normalfordeling ble disse dataene presentert med deskriptiv statistikk i form av median og interkvartilbredde for å vise variasjon i dataene. Disse ble så presentert i en tabell i resultatkapittelet.

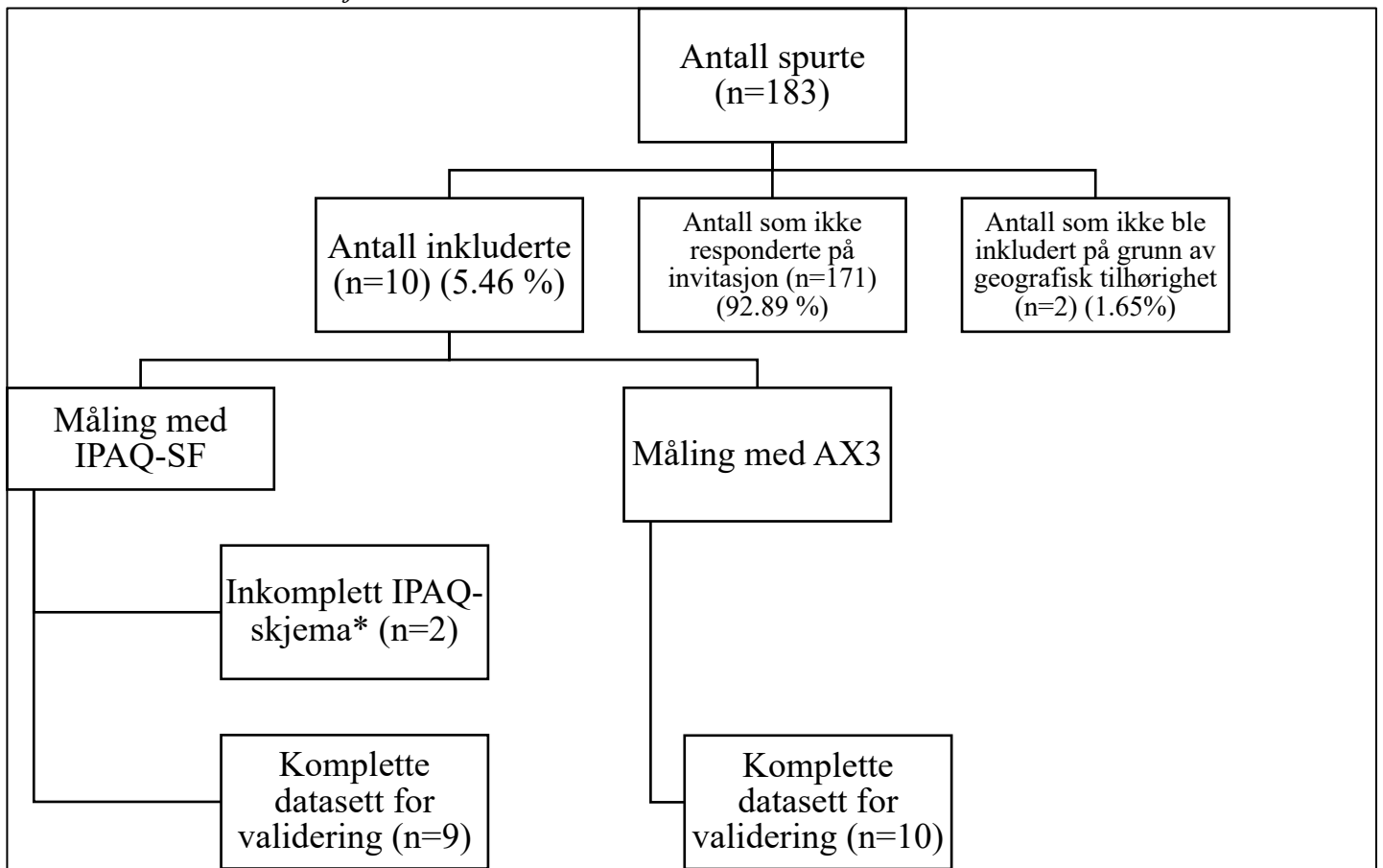
4.0 Resultater

4.1 Utvalg

Av deltakere i BACE-N-studien ble totalt 183 spurt om å delta i dette prosjektet. Totalt 10 deltakere (5.46%) ble inkludert i dette masterprosjektet – i tillegg var det to som responderte positivt på invitasjon til prosjektet som ikke bodde i østlandsområdet (figur 2). Ett skjema måtte ekskluderes fra deler av valideringen av IPAQ-SF grunnet avkrysning på svaralternativet «vet ikke» på moderat fysisk aktivitet (figur 1). Kjønnfordelingen i utvalget var lik mellom kvinner og menn (tabell 1). Median alder for utvalget var 65 år, hvorav 50% av deltakerne var i lønnet arbeid. 30% av deltakerne hadde høyere utdanning enn videregående skole. Ved baselinemåling i BACE-N oppga deltakerne median 5 i smerteintensitet målt med Numeric Rating Scale, til sammenligning var median 1.50 ved måletidspunkt med akselerometer og IPAQ-SF. Fysisk funksjon målt med Roland Morris Questionnaire viste en median score på 5 baseline til BACE-N og ved måletidspunkt til dette prosjektet viste median 0. Måling av SF-36 summert ved fysisk og mental funksjon ved baseline til BACE-N viste skår på henholdsvis median 45.55 og 60.69. Baselinemåling fra BACE-N viste også en skåresum på median 6.5 på Fear Avoidance Beliefs Questionnaire (fysisk aktivitets-komponent). Måling med akselerometer viste at 70, 90 og 100% av utvalget delte inn etter henholdsvis Metode1^{Esliger}, Metode2^{Doherty} og IPAQ-SF, var å regne som tilfredsstillende fysisk aktive etter WHO's retningslinjer.

Figur 1

Flowchart over inklusjon til studien.



*: Ett skjema hvor svaralternativet «vet ikke» ble fylt ut på moderat fysisk aktivitet og ett skjema med «vet ikke» på stillesitting.

Tabell 1
Karakteristika for utvalget

Totalt antall (n=10)	Baseline for BACE-N	Ved måletidspunkt for masterprosjektet
Alder (median (<i>IQR</i>))	65 (6)	
Kjønn , Kvinner (%)	50	
Arbeid:		
Lønnet arbeid (%)	50	
Ikke lønnet arbeid (%)	50	
Utdanning:		
Videregående (%)	70	
Høyere utdanning inntil 4 år (%)	10	
Høyere utdanning 4 eller mer (%)	20	
Smerteintensitet (NRS, range 0-10) (median (<i>IQR</i>))	5 (5)	1.50 (1)
RMDQ (range 0-100 (median) (<i>IQR</i>))	5 (7)	0 (9)
SF-36 (range-100)		
Fysisk funksjon summering (median (<i>IQR</i>))	45.55 (6.65)	
Mental funksjon summering (median (<i>IQR</i>))	60.59 (5.41)	
FABQ – fysisk aktivitets-komponent (range 0-24) (median (<i>IQR</i>))	6.5 (12)	
Fysisk aktivitetsnivå*		
Metode1 ^{Esliger}		(100%)
Metode2 ^{Doherty}		(90%)
Selvrapportert IPAQ-SF (n=9)		(70 %)

NRS= Numeric rating scale, RMDQ=Roland Morris Disability Questionnaire, *andel som tilfredsstillende >150 minutter moderat aktivitet per uke etter inndeling med 10 minuttersintervaller (WHOs anbefalinger). Lav score på NRS, RMDQ og FABQ indikerer god helse. Høy score på SF-36 indikerer god helse.

4.2 Aktivitetsdata fra IPAQ-SF og AX3

Tabell 2

Aktivitetsdata i form av median minutter per uke fra både IPAQ-SF og AX3 akselerometer inndelt etter prosesseringsmetoder.

Intensiteter	IPAQ-SF (n=10*)		AX3 akselerometer	
	Minutter aktivitet per uke	MET-minutter per uke	Metode1 _{Esliger} (n=10) min/uke	Metode2 _{Doherty} (n=10) min/uke
Gange (median (IQR))	780 (503)	2574 (1658)		816 (669)
Lett (median (IQR))			2016 (842)	1528 (625)
Moderat (median (IQR))	420 (690)	1680 (2760)	647 (377)	321 (687)
Meget anstrengende (median (IQR))	165 (439)	1320 (3510)	0 (0)	

*=9 for moderat intensitet

4.3 Samtidig validitet

Tabell 3

Korrelasjonsanalyser mellom IPAQ-SF og AX3 undersøkt med Spearman rang-korrelasjon

IPAQ-SF Intensitet	AX3 Metode1 _{Esliger}	AX3 Metode2 _{Doherty}
Antall deltakere (n)	Spearman korrelasjonskoeffisient τ (p-verdi)	Spearman korrelasjonskoeffisient τ (p-verdi)
Gange (n=10)	-	0.264 (p=0.461)
Moderat (n=9)	-0.233 (p=0.546)	0.733 (p=0.025)
Meget anstrengende (n=10)	-*	-

Signifikante verdier markert med fet skrift. *=Korrelasjon lik 0

Som vist i tabell 3 er det stor variasjon i korrelasjonskoeffisientene på måling av moderat fysisk aktivitet. Metode1_{Esliger} viser en ikke-signifikant korrelasjon på -0.233, mens det for Metode2_{Doherty} er en signifikant korrelasjonskoeffisient på 0.733 som er over den forhåndsdefinerte akseptable korrelasjon på 0.70. For intensiteten meget anstrengende ble det i

etterkant av ekstraksjon av aktivitets-intervaller på >10 minutter på data fra Metode1_{Esliger} kun igjen verdier på 0 minutter som gjør en korrelasjonsanalyse med IPAQ-SF-data uhensiktsmessig.

4.4 Korrelasjon Metode1_{Esliger} og Metode2_{Doherty}

Tabell 2 viser median minutter fysisk aktivitet målt med henholdsvis Metode1_{Esliger} og Metode2_{Doherty} i intensitetene lett og moderat. Resultatene her viser at de to prosesseringsmetodene på akselerasjonsdata fra AX3 fører til ulike resultater uttrykt i minutter aktivitet per uke, noe som også gjenspeiles i Spearman rang-korrelasjon på henholdsvis 0.576 ($p=0.082$) for lett fysisk aktivitet og 0.462 ($p=0.179$) for moderat fysisk aktivitet.

4.5 Feasibility

I målingen på wear-time hos deltakerne viste begge prosesseringsmetodene for validering av wear-time svært like verdier med gjennomsnittlig dager fra 6.94-6.98 med lite spredning oppgitt i interkvartilbredde (tabell 4). Dette ga 10 valide datasett med målinger fra akselerometer med utgangspunkt i lengden på et validmåledøgn på >10 timer og valid uke >4 dager. Sammenligner man datamengden fra AX3 med data fra IPAQ-SF-skjemaene ble ett skjema returnert med respons på «vet ikke» på moderat fysisk aktivitet (figur 1). Fordi moderat fysisk aktivitet er et av tre utfallsmål fra IPAQ-SF måtte derfor ett skjema ekskluderes fra korrelasjonsanalysene på moderat fysisk aktivitet med IPAQ-SF og akselerometer-data.

Tabell 4

Median antall dager med wear time med akselerometre inndelt etter to ulike metoder.

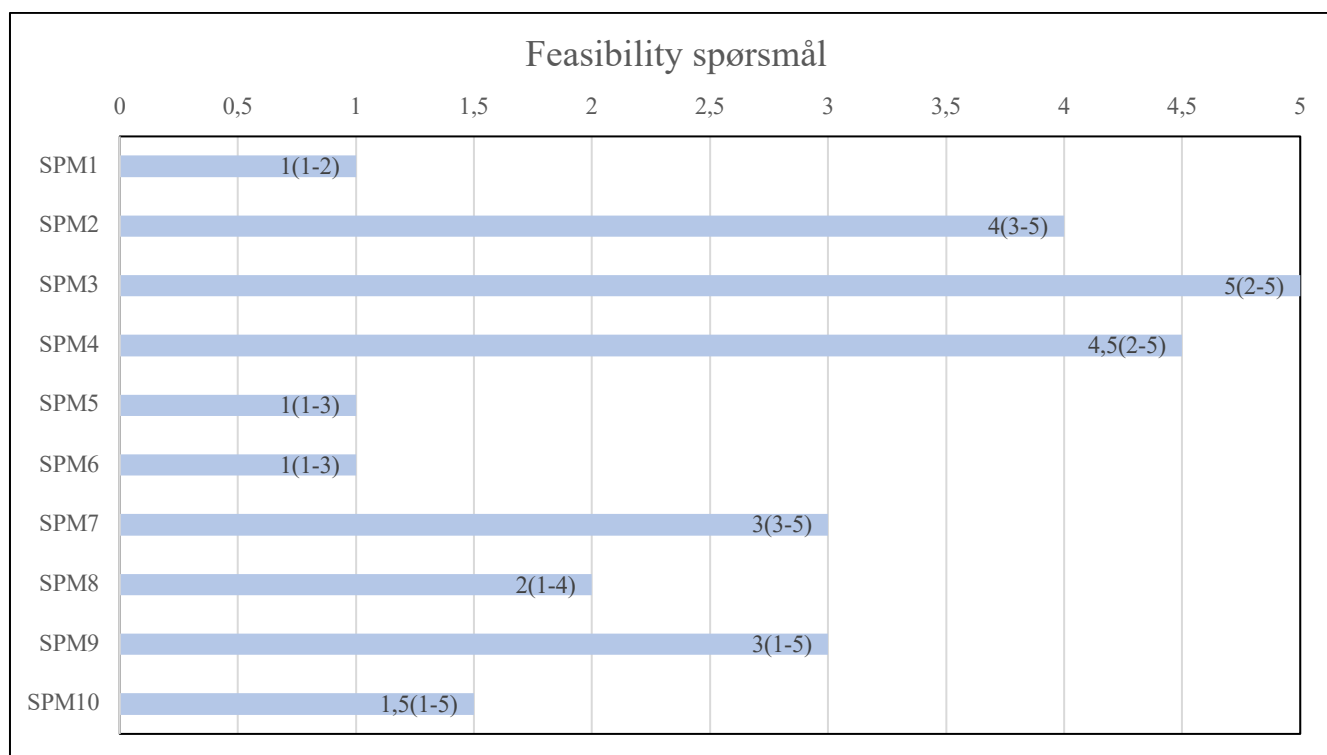
Totalt antall (n=10)	Metode1 _{Esliger} (IQR)	Metode 2 _{Doherty} (IQR)
Median wear-time i dager	6.94 (0.08)	6.98 (0.11)

Spørsmålene som var knyttet til deltakernes erfaring med bruk av IPAQ-SF er angitt i tabell. Minimums- og maksimumsverdiene indikerer at deltakerne var jevnt over enige i svarangivelsen på hver av spørsmålene. Med noen unntak hvor spørsmålene 3, 8, 9 og 10 viste større rekkevidde i minimums- og maksimumsverdier. Disse spørsmålene omhandlet henholdsvis hvorvidt akselerometeret hindret deltakerne i hverdagslige gjøremål, om spørsmålene i IPAQ-SF var forståelige, vanskeligheter med tilbakekalling av egen aktivitet og

om de kunne besvart IPAQ-SF på nytt igjen. Det var en bred enighet knyttet til spørsmål om hvorvidt deltakerne hadde på seg akselerometeret hele tiden underveis i løpet av de totalt 7 døgnene. Spørsmålene som omhandlet reaktivitet og hvorvidt aktiviteten deres ble påvirket av å bære akselerometer, eller svare på IPAQ-SF, viste at de var uenig i at det var noen reaktivitet ved bruk av akselerometer og nøytral til hvorvidt IPAQ-SF ga samme effekt.

Tabell 5

Median svarangivelse for feasibility-spørsmål (vedlegg 2) med minimums og maksimums-verdi angitt i parentes.



5.0 Diskusjon

Dette prosjektet er en metodestudie med hensikt å undersøke den samtidige validiteten til IPAQ-SF med AX3. I tillegg ble aspekter ved feasibility på IPAQ-SF og AX3 akselerometer undersøkt. Deltakernes fysiske aktivitetsnivå, i lys av WHO's offentlige anbefalinger, er også kartlagt. For prosessering av akselerometerdata ble to ulike metoder benyttet i form av en grenseverdi-metode (Metode1_{Esliger}) og en nyere maskinlæringsmetode (Metode2_{Doherty}).

Hovedfunnene i dette prosjektet er høy samtidig validitet ved registrering av selvrapportert moderat fysisk aktivitet når Metode2_{Doherty} benyttes. Ved anvendelse av Metode1_{Esliger} fant man derimot negativ korrelasjon for de samme selvrapporterte dataene på moderat fysisk aktivitet. Det var ingen korrelasjon mellom IPAQ-SF og Metode1_{Esliger} ved måling av meget anstrengende fysisk aktivitet. Selvrapportert tid i gange registrert med IPAQ-SF vurdert med Metode2_{Doherty} viste lav samtidig validitet (tabell 3). Akseptabel samtidig validitet til IPAQ-SF ble dermed funnet kun ved registrering av moderat fysisk aktivitet, og når Metode2_{Doherty} er anvendt (tabell 3). Til slutt viser erfaringer fra deltakere at både utfylling av IPAQ-SF og akselerometer festet til håndleddet er akseptable metoder for måling av aktivitet og at compliance ved bruk av akselerometer på håndleddet var høy. Majoriteten av deltakerne oppfyller kravene til fysisk aktivitet fra WHO, men oppfylleelsesgraden varierer noe avhengig av målemetode. Korrelasjonsanalyser mellom de to ulike prosesseringsmetodene benyttet i dette prosjektet viste ikke-signifikant lav til moderat korrelasjon, noe som indikerer en diskrepans mellom de ulike prosesseringsmetodene. Diskrepansen tyder på at de to prosesseringsmetodene klassifiserer fysisk aktivitet fra rådata ulikt, noe som kan delvis forklare ulike funn på samtidig validitet ved måling av moderat fysisk mellom IPAQ-SF og de to prosesseringsmetodene.

5.1 Resultater

5.1.1 Samtidig validitet

Resultatene av korrelasjonsanalysene som fremkommer i denne oppgaven er primært preget av et lite utvalg, noe som fører til to ting. Det ene er at resultatene kan skyldes tilfeldigheter. Det andre er at et begrenset utvalg generelt sett er å anse som mindre representativt enn et større, noe som medfører redusert grad av ekstern validitet i dette masterprosjektet (Carter & Lubinsky, 2016, s. 103). I tillegg er en begrensning den manuelle behandlingen av data for

inndeling av 10-minutters intervaller som kan ha gitt tapte data. Det er også noe usikkerhet knyttet til AX3 som kriteriestandard. Disse feilkildene er essensiell i tolkningene av resultatene som fremkommer på undersøkelse av samtidig validitet til IPAQ-SF med AX3 i dette masterprosjektet.

For vurdering av samtidig validitet til gange målt med IPAQ-SF med Metode2^{Doherty} viste korrelasjonsanalysene en svak ikke-signifikant korrelasjon på 0.264 (tabell 3). Dette er i samsvar med studien av Grimm et al. (2012) som undersøkte gange målt med IPAQ-SF versus moderat gange fra akselerometer (0.276). Det samme resultatet viser studien av Hurtig-Wennlof et al. (2010) hvor det var lav korrelasjon (0.297). Korrelasjonen på 0.264 som fremkommer i dette prosjektet er derfor å regne som lav, og under det som er satt som akseptabel verdi i en valideringsstudie (Terwee et al., 2007). Her er en mulig årsaksforklaring på den lave korrelasjonen at deltakerne har rapportert gående som har vart i underkant av 10 minutter. I studien Copeland and Esliger (2009) på 38 friske eldre viste at 66% av aktiviteten, hvorav mesteparten av aktivitet var gange, varte i underkant av 10 minutter. Det er derfor tenkelig at det er en sannsynlighet for at deltakerne har overrapportert sekvenser av gange som lengre enn de faktisk var.

Validering av moderat fysisk aktivitet med IPAQ-SF i dette prosjektet resulterte i Spearman korrelasjonskoeffisient på 0.744 sammenlignet med Metode2^{Doherty} (tabell 3). Dette er å regne som akseptabel validitet (Terwee et al., 2007). Det er også høyere enn hva andre valideringsstudier har funnet. Studien av Grimm et al. (2012) viste imidlertid en korrelasjon på 0.042 ved måling av moderat fysisk aktivitet. Studien av Grimm et al. (2012) har noen ulikheter med masterprosjektet, da annet ActiGraph ble benyttet. Lav korrelasjon ble imidlertid også funnet i den systematiske oversiktsstudien av P. H. Lee et al. (2011) hvor det oppgis korrelasjon på 0.37 mellom akselerometer og IPAQ-SF ved måling av moderat fysisk aktivitet. Utvalget hos og akselerometer var imidlertid forskjellig fra utvalg og akselerometer i dette masterprosjektet. I en nyere studie av C. Cleland, Ferguson, Ellis, and Hunter (2018) hvor de gjennomførte validering av IPAQ Long form på et utvalg eldre i Storbritannia viste korrelasjonskoeffisienten en verdi på 0.577 ved måling av moderat-til-anstrengende fysisk aktivitet. Metodisk er studien av C. Cleland et al. (2018) ulik dette masterprosjektet ved at IPAQ Long form ble benyttet, samt antall minutter moderat og meget anstrengende aktivitet ble summert sammen til en intensitet (moderat-til-anstrengende). IPAQ Long form skiller seg en del fra IPAQ-SF ved at man har delt inn i fysisk aktivitet i transport, fritid og arbeid. Måten

spørsmålene stilles er likevel lik spørsmålene i IPAQ-SF og har derfor en viss overføringsverdi. En annen studie av Hurtig-Wennlof et al. (2010) på 54 eldre viste en korrelasjonskoeffisient på 0.396 mellom selvrapportert moderat fysisk aktivitet fra IPAQ-Elderly og akselerometer. IPAQ-Elderly avviker noe fra IPAQ-SF ved at man har omrokkert på rekkefølgen til spørsmålene på gange, moderat og meget anstrengende. På grunnlag av at man i dette prosjektet har funnet en høyere korrelasjon på måling av moderat aktivitet sammenlignet med andre studier, er det vanskelig å konkludere med en spesifikk årsaksforklaring på hvorfor man finner en høyere korrelasjon ved måling av moderat fysisk aktivitet. Man kan derfor ikke utelukke at disse resultatene skyldes tilfeldigheter.

I motsetning til resultatene på korrelasjon mellom IPAQ-SF og Metode2^{Doherty} ved måling av moderat fysisk aktivitet, ses det en ikke-signifikant negativ korrelasjon mellom Metode1^{Esliger} og IPAQ-SF (tabell 3). Dette kan tyde på en diskrepans mellom konstruktet (moderat fysisk aktivitet) som IPAQ-SF og Metode1^{Esliger} avspeiler, en hypotese som støttes av at aktivitetsdata fra Metode1^{Esliger} viser at median minutter moderat fysisk aktivitet er høyere enn det som har blitt rapportert gjennom IPAQ-SF (tabell 2). En mulig forklaring på den negative korrelasjonen er hvorvidt prosessering av akselerometerdata med Metode1^{Esliger} har ført til en overestimering av antall minutter moderat aktivitet ved måling med Metode1^{Esliger}. Det kan derfor tyde på at det er en overlapp i grenseverdiene benyttet i Metode1^{Esliger} (4.00-6.99 MET) for moderat fysisk aktivitet og at minutter tilbragt i andre intensiteter kan ha blitt oppfattet som moderate.

For utvalget i dette prosjektet var det en overrapportering av meget anstrengende fysisk aktivitet ved måling av IPAQ-SF hvor selvrapportert median minutter per uke viste 165 minutter, til sammenligning med 0 minutter målt med akselerometer og Metode1^{Esliger} (tabell 2). Mangelen på korrelasjon ved måling av meget anstrengende fysisk aktivitet er i rimelig samsvar med hva som ble funnet i studien av Grimm et al. (2012), hvor korrelasjonen var ikke-signifikant verdi på 0.166. Det interessante i studien av Grimm et al. (2012) var at korrelasjonstallene på aktivitetsdata fra akselerometeret var ulike etter inndeling av ekstraherte intervaller på over 10 minutter og de aktivitetsdataene hvor ekstraksjon ikke var gjennomført. I forkant av ekstraksjon på aktivitetsdataene i dette masterprosjektet viste aktivitetsdataene at deltakerne hadde tilbragt et minimalt antall minutter i meget anstrengende fysisk aktivitet, men etter ekstraksjon av 10 minutters intervaller viste dataene 0 minutter. En mulig hypotese på resultatene fra måling av meget anstrengende aktivitet er at antall rapporterte minutter meget anstrengende aktivitet fra deltakerne har blitt registrert av Metode1^{Esliger} som moderat aktivitet istedenfor meget

anstrengende. Dette kan tenkes å være sannsynlig siden man så en overestimering av minutter moderat fysisk aktivitet målt med Metode1_{Esliger} i forhold til minutter moderat aktivitet målt med IPAQ-SF. I tillegg kan det tenkes at begrepet meget anstrengende muligens kan ha blitt misforstått av deltakerne, hvor aktiviteten de tenker har vært meget anstrengende ikke var i tråd med intensiteten målt av Metode1_{Esliger} (Bahr & Helsedirektoratet, 2009, s. 127).

Oppsummert for resultatene på beregning av samtidig validitet til IPAQ-SF viser at man fant tilfredsstillende samtidig validitet ved måling av moderat fysisk aktivitet vurdert med Metode2_{Doherty}. Funnet i dette masterprosjektet viser også en høyere validitet enn hva andre sammenlignbare studier har funnet. Derimot fant man negativ korrelasjon mellom IPAQ-SF og Metode1_{Esliger} ved måling av moderat aktivitet. For gange og meget anstrengende fant man henholdsvis lav til ingen korrelasjon, noe som er mer i samsvar med hva andre studier har funnet. Resultatene av dette masterprosjektet må imidlertid sees i lys av det lave antall inkluderte (n=10) og den manuelle behandlingen av aktivitetsdata fra akselerometer hvor man hentet ut 10-minutters intervaller. Som det blir diskutert videre nedenfor i diskusjonen kan dette ha vært en feilkilde i form av tapte eller feilberegning av aktivitetsdata fra akselerometer.

5.1.2 Korrelasjon mellom Metode1_{Esliger} og Metode2_{Doherty}

Korrelasjonsanalysene mellom Metode1_{Esliger} og Metode2_{Doherty} viste begge ikke-signifikante verdier på 0.576 og 0.462 ved måling av henholdsvis lett og moderat fysisk aktivitet (punkt 4.4). Variasjonen man ser mellom disse to prosesseringsmetodene er i tråd med andre lignende studier hvor man har benyttet to ulike prosesseringsmetoder for å kvantifisere og klassifisere rådata fra akselerometer (Kerr et al., 2017; Van Holle, De Bourdeaudhuij, Deforche, Van Cauwenberg, & Van Dyck, 2015).

Et annet aspekt av resultatene fra korrelasjonsanalysene på prosesseringsmetodene er som det er påpekt i den systematiske oversiktsartikkelen av P. H. Lee et al. (2011) hvor majoriteten av valideringsstudiene på IPAQ-SF gjennomført med ulike akselerometer og prosesseringsmetoder viser stor variasjon. Den samme variasjonen i korrelasjon mellom IPAQ-SF og akselerometer som er rapportert av P. H. Lee et al. (2011) finner man også igjen i resultatene fra dette prosjektet (punkt 4.4).

Korrelasjonen på lett fysisk aktivitet målt med de to ulike prosesseringsmetodene kunne tenkes å være svakere dersom man ikke hadde slått sammen minutter aktivitet fra lett og gange fra

Metode₂^{Doherty}. Dette ble likevel gjennomført fordi man hadde mistanke om at lett fysisk aktivitet målt med Metode₁^{Esliger} også omfattet gange. For å styrke grunnlaget for sammenligning mellom de to metodene var dette et tiltak som antakeligvis har høynet korrelasjonen mellom de to ulike prosesseringsmetodene.

5.1.3 Feasibility

Resultatene fra undersøkelse av feasibility viste at man oppnådde et høyt antall valide måledøgn med begge prosesseringsmetoder (tabell 4). Jevnt over var deltakerne positive til bruk av både IPAQ-SF og AX3 Akselerometer. I vurderingen av feasibility ved bruk av både IPAQ-SF og AX3 akselerometer ble det i teori-kapittelet (punkt 2.3.2) belyst ulike begreper innenfor feasibility som danner grunnlaget for hvordan vurderingen av feasibility blir gjennomført. Nedenfor følger diskusjonsmomenter knyttet til resultatene på wear-time valideringen, spørreskjemaet på feasibility og noen andre momenter knyttet til bruken av IPAQ-SF og AX3 akselerometer. En betydelig begrensning på som må tas i betraktning på resultatene fra spørreskjemaet knyttet til feasibility var at dette var spørsmål som var adapterte fra andre spørreskjemaer. Spørreskjemaet i sin helhet har derfor ikke vært utprøvd i tidligere studier.

Et viktig funn i dette prosjektet var antall valide måledøgn med akselerometer (tabell 4). Her var resultatene mellom de ulike prosesseringsmetodene like og viste at alle deltakerne hadde 7 valide måledøgn, til tross for at Metode₁^{Esliger} og Metode₂^{Doherty} hadde ulike utgangspunkt for validering av wear-time. Til sammenligning med resultatene fra studien av Hurtig-Wennlof et al. (2010) hvor 20 % av tapte data kunne tilegnes akselerometer-data. I studien av Hurtig-Wennlof et al. (2010) ble akselerometeret festet til livet med et elastisk band, i motsetning til dette prosjektet hvor håndleddet ble benyttet. Dette kan tyde på at håndleddet er en velegnet anatomisk struktur for måling av fysisk aktivitet over lengre tid. Dette ble også bekreftet av erfaringene til deltakerne hvor det er klare indikasjoner på at det var lite ubehag knyttet til å ha akselerometeret på håndleddet (tabell 4).

Når det gjelder anvendbarhet av IPAQ-SF var utfordringer, sett fra en forsker/klinikers perspektiv, at svaralternativet «vet ikke» byr på utfordringer. Dette viste seg i form av to ukomplette IPAQ-SF-skjemaer (figur 1). Til tross for at man ved inklusjonsmøte ga instruksjon om å ikke svare «vet ikke», viser dette at det er en potensiell risiko for tapte data ved bruk IPAQ-SF. Noe som også var gjeldende i valideringsstudien av Sanda et al. (2017), hvor 3 skjema ble returnert ukomplett. Fitzpatrick et al. (1998) argumenterer for at dersom en deltaker

returnerer tilbake et ukomplett skjema, kan dette være indikasjoner på at spørreskjemaet er vanskelig å forstå. Her kommer man litt tilbake til diskusjonen om hvorvidt begrepsbruken i IPAQ-SF gir rom for individuell tolkning. Som en del av spørreskjemaet på feasibility var spørsmål 8 direkte knyttet til opp mot hvordan deltakerne oppfattet spørsmålene i IPAQ-SF (vedlegg 1). Median svarangivelse ga antydning til at deltakerne var enige i at det var lett å forstå disse spørsmålene, mens variasjonen målt med minimums og maksimums-verdi indikerte stor variasjon i denne oppfatningen. Dette kan igjen tyde på at enkelte av deltakerne synes spørsmålene var vanskelige. Resultatene tatt i betraktning så tyder det på at AX3 har en fordel i form av valide måledata, hvor man opplevde to ukomplette IPAQ-SF.

Siden ingen av deltakerne vil bli direkte berørt av i hvilken grad Metode1_{Esliger} eller Metode2_{Doherty} benyttes, vil dette momentet holdes utenfor vurderingen av feasibility. Det blir derfor fokus på i hvilken grad dette har noe å si for forskeren/undersøkeren. I vurderingen av implementering av Metode1_{Esliger} ovenfor forskeren/undersøkeren er dette en gjennomførbar metode som krever lite forkunnskap av den som skal gjennomføre prosesseringsmetoden. Det krever tilgang på datamaskin med Windows-programvare og OMGUI som lastes ned gratis. Når data er ferdig innsamlet er det innebygd i OMGUI en rekke ulike algoritmer med ulike grenseverdier basert på ulike aldersgrupper som gir antall minutter i ulike intensiteter. Sett fra et implementerings-perspektiv er det lite som tilsier at Metode1_{Esliger} ikke skal kunne inngå i en klinisk hverdag hos fysioterapeut. Derimot i vurdering av implementering med Metode2_{Doherty} var det knyttet utfordringer i form av at programvaren kunne utelukkende benyttes på datamaskin med Linux. I tillegg var det utfordringer knyttet til installering og bruk av programvaren siden det var her snakk om kommando-basert programvare. Prosessering av rådata med Metode2_{Doherty} kan derfor regnes som mer utfordrende å gjennomføre enn i forhold til prosessering med Metode1_{Esliger}.

5.1.4 Statistiske valg

På grunn av et lite utvalg og data som ikke var normalfordelte, ble det benyttet en Spearman rang-analyse som statistisk metode for å undersøke relasjon mellom variablene. I dette masterprosjektet er det benyttet ulike prosesseringsmetoder for å beregne fysisk aktivitet på, og man ser at det er stor variasjon i relasjonen mellom akselerometerdata og IPAQ-data. I teorien har man med en Spearman rang-analyse utelukkende målt den monotone relasjonen mellom de ulike metodene hvor man forventer at dersom variabel X øker i verdi, så skal også variabel Y øke. En alternativ måte for å vise et noe mer nyansert bilde enn hva man får frem fra en

korrelasjonsanalyse er å utforme Bland & Altman-plots hvor man plotter inn differansen og gjennomsnittet til de to ulike variablene man vil vurdere inn i et scatter-plot. Videre beregner man øvre og nedre del av enigheten mellom variablene, og man får dermed en mulighet til å vurdere datapunktene og hvorvidt disse holder seg innenfor disse grensene for enighet mellom de to variablene (Giavarina, 2015). Det ble forsøkt gjennomført en Bland & Altman-analyse for dataene i dette masterprosjektet, men grunnet det lave utvalget var det lite hensiktsmessig å ta med dette plottet inn i resultat-kapittelet siden den begrensede mengden data ga usikkert grunnlag for tolkning.

5.2 Metode

5.2.1 Design

For å kunne gjennomføre dette masterprosjektet ble et tverrsnitts-design valgt. Dette designet er også det som egner seg for å kunne svare problemstillingen knyttet til undersøkelse av samtidig validitet til IPAQ-SF med AX3 akselerometer (Kesmodel, 2018). Designet muliggjorde at man i tillegg til å undersøke den samtidige validiteten til IPAQ-SF kunne vurdere forhold knyttet til feasibility til både IPAQ-SF og AX3 akselerometer.

Et tverrsnittsdesign vil si at man innhenter data kun ved en anledning (Kesmodel, 2018). I dette tilfellet ble fysisk aktivitet hos deltakerne ble målt ved en anledning (sammenhengende over totalt 7 døgn). Siden et tverrsnittsdesign gir kun et øyeblikksbilde av det totale aktivitetsnivået til deltakerne i dette prosjektet og man har derfor ikke belegg for å kunne si noe om endring i aktivitetsnivået til deltakerne over tid. Eksempelvis så kan ikke dataene som fremkom i dette prosjektet si noe om årsaksforhold knyttet til hvorfor aktivitetsnivået var som det var, og man kan heller ikke si noe konkret om hvorvidt deres ryggmerter kan ha påvirket aktiviteten deres. Det er heller ingen mulighet for å kunne si noe om årsaksforhold knyttet til bias ved selvrapporingen av fysisk aktivitet, hvor evidens fra teori viste at faktorer som alder, grad av ryggmerter og utdanningsnivå har en innvirkning på bias ved selvrapporing (punkt 2.2.2). En ytterligere utvidelse av studiet kunne derfor ha vært en mer grundig analyse på om hvorvidt høyere alder, ryggmerter og utdanningsnivå har påvirket selvrapportert fysisk aktivitet.

5.2.2 Rekruttering og utvalg

At man i dette masterprosjektet har rekruttert deltakere fra BACE-N har noen utfordringer i form av hvilke deltakere, og ikke minst hvor mange deltakere som ble inkludert i dette prosjektet. Det faktum at man rekrutterte utelukkende BACE-N-deltakere til dette masterprosjektet har lagt sterke føringer for hvilke deltakere som ble inkludert i dette prosjektet og kan muligens ha bidratt til en seleksjonsskjevhet i utvalget. Nedenfor diskuteres mulig seleksjonsskjevhet i utvalget.

Ett av inklusjonskriteriene for BACE-N var blant annet at deltakere skulle ha 6 måneder uten konsultasjon hos behandlere for rygg smerter i forkant av besøk hos primærhelsetjenesten med en ny episode med rygg smerter. Dette kan tenkes å ha ekskludert en del deltakere med ryggplager som kunne ha bidratt til et mer variert utvalg. En ytterligere faktor som kan ha bidratt til å redusere graden av ekstern validitet var det faktum at kun deltakere som var under oppfølging i BACE-N kunne rekrutteres til masterprosjektet og ikke nyrekrutterte deltakere fra BACE-N. Som følger av at deltakerne i BACE-N allerede var under oppfølging kan dette tenkes å ha påført dem tretthet og mindre motivasjon til å bidra til annen forskning.

En sannsynlig konsekvens av det smale utvalget i dette prosjektet er at man oppnår en lav grad av ekstern validitet på resultatene fra dette masterprosjektet (Carter & Lubinsky, 2016, s. 103). Et ytterligere inklusjonskriterie som måtte tillegges dette prosjektet var at deltakerne måtte befinne seg på Østlands-området. Dette ble lagt til for å gjøre prosjektet praktisk gjennomførbart for masterstudenten. To deltakere ble derfor ekskludert på grunn av for lang reiseavstand. Den geografiske begrensningen kan derfor også ha bidratt til et ytterligere smalere utvalg med tanke på geografisk variasjon.

Utvalget på totalt 10 deltakere er å regne som lite. Responsraten på 5% (figur 1) som fremkommer i resultatene viser at rekrutteringen til studiet ikke gikk som planlagt. Den lave responsraten kan imidlertid tenkes å ikke være unikt for dette prosjektet, da man ser tendenser til lav responsrate i Europa (Mindell et al., 2015; Williams, Entwistle, Haddow, & Wells, 2008). Dette gir derfor noe grunn til å tro at den lave responsraten ikke er eksklusivt for det utvalget som ble invitert.

Som nevnt i metoden (punkt 3.1) var årsaken til at man benyttet seg av deltakere fra BACE-N på grunnlag av ønske fra prosjektledelsen om å undersøke samtidig validitet til IPAQ-SF med AX3 akselerometer. De nevnte ulempene ved å rekruttere deltakere via BACE-N-studien kunne imidlertid vært unngått ved å rekruttere direkte fra primærhelsetjenesten. Dette ville imidlertid krevd en helt egen søknad til REK, noe som igjen ville betydd måneder med forsinket oppstart av masterprosjektet i påvente av godkjenning. Det var derfor ikke anledning til å kunne gjennomføre dette. Til slutt er det viktig å nevne at det var og noen deltakere som ved inklusjon til masterprosjektet trodde at de skulle ha et ytterligere møte med masterstudenten for å avlevere akselerometeret. Det er derfor mulig at informasjonen i invitasjonen til masterprosjektet ikke var tydelig nok på at akselerometeret skulle returneres i en ferdigfrankert konvolutt.

Utvalgets karakteristika er beskrevet i tabell 1 i resultatkapittelet. Her fremkommer det blant annet at deltakernes aktivitetsnivå målt med akselerometer viste at majoriteten av utvalget kunne betegnes som tilfredsstillende fysisk aktive jamfør WHO's retningslinjer. Årsaken til at man kun hadde moderat aktivitet som utfallsmål på om de var aktive eller ikke var på grunnlag av at ingen av deltakerne oppnådde 75 minutter meget anstrengende fysisk aktivitet per uke målt med akselerometer. Til sammenligning med nasjonale data viser den siste folkehelse rapporten fra 2015 at blant 3020 individer i aldersgruppen 50-64 år er 37.1% regnet som fysisk aktive, mens i aldersgruppen >65 år er det et noe lavere antall (31.5 %) som er fysisk aktive (Helsedirektoratet, 2015). En viktig bemerkning ved sammenligning med utvalget i folkehelse rapporten er at de i denne rapporten har slått sammen moderat og anstrengende aktivitet, til forskjell fra dette prosjektet hvor man kun fikk valide data på moderat fysisk aktivitet. Dette kan delvis forklare de ulike data som har fremkommet i dette prosjektet, sammenlignet med folkehelse rapportens data. Det er derimot viktig å ta i betraktning det lave antall deltakere i dette prosjektet og at dette i seg selv gir et svakt grunnlag for å sammenligne resultater fra dette prosjektet med utvalget fra den Norske folkehelse rapporten.

Hvorvidt deltakerne var tilfredsstillende fysisk aktive ble også beregnet ved å bruke selvrapporterte data fra IPAQ-SF, hvor man grunnet avkryssing for svaralternativet «vet ikke» i ett skjema hadde valide data fra kun 9 deltakere (tabell 1). Her viste resultatene på fysisk aktivitet rapportert via IPAQ-SF at 70 % kunne betegnes som fysisk aktive (tabell 1). En tverrsnittundersøkelse på sammenligning av BACE-deltakere fra Brasil og Nederland brukte den kategoriske variabelen for scoring av IPAQ-SF for å klassifisere aktivitetsnivået til deltakerne (F. R. Jesus-Moraleida et al., 2017). Den kategoriske inndelingen gir kategorier som

lavt, moderat og høyt fysisk aktivitetsnivå, hvor moderat fysisk aktivitetsnivå indikerer mindre enn 60 minutter anstrengende aktivitet i uka, mindre enn 150 minutter moderat og/eller gange, samt en totalsum på 600 MET-minutter (www.ipaq.ki.se). I studien av F. R. Jesus-Moraleida et al. (2017) viste resultatene at 62% av de Brasilianske deltakerne kunne kategoriseres lavt fysisk aktive, sammenlignet med de Nederlandske deltakerne hvor 26% hadde lavt aktivitetsnivå. Primært viser dette en stor variasjon i selvrapporingen mellom de to ulike utvalgene fra Brasil og Nederland. Sekundært viser resultatene fra Nederland en relativ lik rapportering av fysisk aktivitet som fra utvalget i dette prosjektet hvor 70% av deltakerne kunne defineres som fysisk aktive i dette prosjektet, sammenlignet med omtrent 70% i det Nederlandske utvalget. Vurdering av hvorvidt deltakerne var fysisk aktive med selvrapporterte data ble i dette prosjektet gjort på grunnlag av 9 deltakere og resultatene må derfor tolkes med forsiktighet. Aktivitetsnivået som ble målt i dette prosjektet var likevel markant høyt, noe som kan være en konsekvens av at man har tiltrukket seg deltakere som er å regne som mer aktive enn gjennomsnittet. Sådan kan dette tale for en seleksjonsskjevhet i utvalget.

Grad av smerteintensitet målt med Numeric Rating Scale (NRS) viser en median nedgang fra måling ved inklusjon BACE-N 5 (IQR:5) til 1.50 (IQR:1) målt ved inklusjon i dette masterprosjektet (tabell 1). I en studie på 2854 pasienter med kronisk muskel- og skjelettsmerte av Boonstra et al. (2016) ble cut-off verdier på hvilke NRS-verdier som kan betegnes som mild (>5), moderat (6-7) og betydelig smerte (>8) foreslått. Med tanke på at NRS-score for utvalget i dette prosjektet ligger betydelig under den foreslåtte verdien på >5 i NRS, indikerer dette at deltakerne i dette prosjektet ikke lenger var like plaget av sine rygg smerter som ved inklusjon til BACE-N. Ulikhetene på median smerteintensitet som fremkommer i dette prosjektet kan være en av flere mulige årsaksforklaringer på hvorfor de i utgangspunktet takket ja til å delta i dette masterprosjektet. Den lave graden av smerteintensitet hos dette utvalget kan muligens ses i sammenheng med det høye aktivitetsnivået til deltakerne hvor man fant at 90-100 % av deltakerne var å regne som fysisk aktive (tabell 1). Til tross for sparsom evidens, så viser en studie på 51 eldre, hvor fysisk aktivitet ble målt med akselerometer, at mindre sedat tid og økt grad av aktivitet på fritiden gir økt smerteinhibisjon (Naugle, Ohlman, Naugle, Riley, & Keith, 2017). Dette støttes også opp av tidligere studie som har vist at ulike former for fysisk aktivitet kan ha en effekt på smerter (Geneen et al., 2017). Variablene på aktivitetsnivået og smerte presentert i tabell 1 kan derfor tenkes å ses i sammenheng med hverandre.

Parallelt med nedgangen i median smerteintensitet målt med NRS, var det også en endring i score på Roland Morris Disability Questionnaire på baseline-data sammenlignet med score ved inklusjon til dette prosjektet (tabell 1). Dette gir ytterligere grunnlag for å hevde at et flertall av deltakerne i dette prosjektet opplevde å være i mindre grad plaget av ryggsmarter underveis i målingene med både IPAQ-SF og AX3 akselerometer.

Data på livskvalitet målt med SF-36 viser at utvalget scoret median 45.55 og 60.59 for henholdsvis de fysiske og mentale komponentene ved inklusjon til BACE-N (tabell 1). Til sammenligning viser tallene fra måling på BACE-deltakere fra Nederland gjennomsnittsverdier på 43.20 og 49.56 på henholdsvis de fysiske og mentale komponentene, sammenlignet med 41.65 og 42.30 fra Brasilianske deltakere (F. R. D. Jesus-Moraleida et al., 2018). Her viser skårsummene på den fysiske komponenten av SF-36 at utvalget i dette masterprosjektet skårer relativt likt med utvalgene fra Nederland og Brasil, mens for den mentale komponenten skårer utvalget i dette prosjektet noe høyere – som indikerer en mer positiv vurdering av egen livskvalitet.

Grad av fear-avoidance målt med Fear-avoidance Beliefs Questionnaire for fysisk aktivitet ved inklusjon til BACE-N viste for dette utvalget en median score på 6.5 (tabell 1). Sammenligning med andre BACE-utvalg fra Brasil og Nederland viser gjennomsnittlige verdier på henholdsvis 15.77 og 13.43, noe som indikerer at utvalget i dette prosjektet med median 6.5 skårer betraktelig lavere enn utvalgene i Brasil og Nederland (F. R. D. Jesus-Moraleida et al., 2018). Dette kan indikere at utvalget i dette masterprosjektet var en seleksjon av deltakere som var fortrolig med fysisk aktivitet, til tross for ryggsmarter. Her er det viktig å bemerke at man kun har benyttet seg av første del av Fear-avoidance Beliefs Questionnaire i dette masterprosjektet, som består av totalt fem spørsmål knyttet til fysisk aktivitet og fear-avoidance. Et annet viktig aspekt ved dataene fra Fear-avoidance Beliefs Questionnaire og SF-36 er at begge to kun ble besvart ved inklusjon til BACE-N, og det kan derfor ikke utelukkes at skårsummene kan ha endret seg når deltakerne ble inkludert i dette masterprosjektet, slik som man så ved endringene i NRS og RMDQ.

Oppsummert viser dataene at dette var å regne som et aktivt utvalg, med liten grad av smerteintensitet og nedsatt funksjon. Derfor er det stor usikkerhet knyttet til hvorvidt utvalget i dette masterprosjektet er et representativt utvalg for populasjonen av mennesker over 55 år med ryggsmarter.

5.2.3 Etikk

En etisk overveielse som er viktig å ta ved gjennomførelse av studier er hvorvidt intervensjonen eller målingen er å regne som en byrde for deltakerne (Carter & Lubinsky, 2016, s. 44). Tiden på hvor lang tid hver av deltakerne måtte sette av for inklusjon har for dette prosjektet blitt holdt på et lavt nivå. Samtlige deltakere ble oppsøkt på et sted hvor det var enkelt for dem å treffes. Utfylling og postlegging av IPAQ-SF-skjemaet er også vurdert i form av tid, hvor selve utfyllingen av IPAQ-SF ikke er beregnet til mer enn 10-15 minutter. Imidlertid kan postlegging kan være noe mer tidkrevende. En forbedring her kunne vært å gjennomføre utfylling av IPAQ-SF elektronisk.

Et annet aspekt er vurdering av byrden med måling med AX3 akselerometer. Siden man i dette prosjektet har innhentet opplysninger om erfaringene som deltakerne gjorde seg i løpet av måletiden (diskuteres nærmere i punkt 5.2.3) gir dette et grunnlag for å vurdere byrden som ble påført deltakerne underveis i prosjektet. Ifølge produsenten av klokke-reimen til AX3 akselerometer er selve reimen laget av et godkjent silikon-materiale som skal unngå å irritere huden, og låsemekanismen er laget av rustfritt stål (www.axivity.com). Dette gjør at man har minimert risikoen for eventuelle hudplager hos deltakerne. I tillegg rapporterte deltakerne at akselerometeret ikke var til hinder for dem i hverdagslige gjøremål (tabell 5).

Masterprosjektet var videre godkjent av NSD ved egen endringsmelding, og alle deltakerne undertegnet informert skriftlig samtykke før studiespesifikke prosedyrer ble iverksatt. Til tross for nøytral informasjon i invitasjonen til å delta i dette prosjektet, kan man ikke utelukke at noen følte press til å delta i prosjektet. Rollen som både undersøker og forfatter av dette prosjektet kan ha bidratt til økt grad av bias i form av social desirability hos deltakerne i selvrapporeringen av fysisk aktivitet. Social desirability kan kort defineres som tendensen deltakere i studier har til å overrapportere ønsket adferd – eksempelvis fysisk aktivitet og sunt kosthold. Uønsket adferd, som eksempel konsum av alkohol og sedat tid, tenkes å underrapporteres (Paulhus, 1991).

5.2.4 Aktivitetsdata fra rapportering med IPAQ-SF

Tidligere i oppgaven (punkt 2.2.3) ble studier på validitet og reliabilitet til IPAQ-SF redegjort for, som tyder på at validiteten til IPAQ-SF er lav. Delvis kan dette forklares med ulike metodiske valg i valideringsstudiene (P. H. Lee et al., 2011). Man kan likevel ikke komme

utenom at majoriteten av valideringsstudiene peker i retning av at IPAQ-SF har en lav validitet. I studien av Craig et al. (2003) som gikk på tvers av 14 ulike land fant en lav validitet og høy reliabilitet av IPAQ-SF. Det er imidlertid ikke anbefalt bruk av Spearman rang-korrelasjon ved reliabilitetstesting (Terwee et al., 2007).

Ved inklusjon til masterprosjektet ble deltakerne grundig instruert i å ikke svare «vet ikke» på spørsmålene i IPAQ-SF. Slike instruksjoner avviker fra hvordan man vanligvis administrerer IPAQ-SF ved at man skal la respondenten lese og bruke IPAQ-SF slik som det står beskrevet i skjemaet. Det ble likevel gjort som preventivt tiltak for å hindre ukomplette skjema, siden det er evidens på at registrering av aktivitetsdata ved bruk av IPAQ-SF kan gi en del tapte data (Sanda et al., 2017). Det kan tenkes ha bidratt til påvirkning av responsen fra deltakerne på de spørsmålene hvor de normalt sett ville svart «vet ikke», til at de gir mer usikre estimater på deres egen fysiske aktivitet.

IPAQ-SF inngår som spørreskjema på fysisk aktivitet i alle oppfølgingstidspunktene BACE-N. Det vil si på baseline, 3, 6, 12 og 24 måneders oppfølging. Deltakerne fylte også ut spørreskjemaet to ganger i løpet av dette masterprosjektet, hvor det opprinnelig var planlagt en undersøkelse av test-retest reliabilitet av IPAQ-SF i tillegg til vurdering av samtidig validitet. Det faktum at deltakere i dette prosjektet har fylt ut IPAQ-SF i såpass utstrakt grad gjennom BACE-N, samt dette masterprosjektet, kan ha ført til en viss form for innlæring og kjennskap til skjemaet. Selv om det, ut ifra det undertegnede har greid å bringe på det rene, ikke foreligger evidens på hvorvidt innlæring av IPAQ-SF kan ha en effekt på hvordan respondenter rapporterer fysisk aktivitet, så kan man ikke utelukke at deltakerne gjennom gjentatte utfyllinger av IPAQ-SF kan ha fått mer kjennskap til skjemaet og ikke minst et mer bevisst forhold til eget aktivitetsnivå. Dette kan på en annen side ha vært med på å minske effekten av recall bias, som kan kort kan defineres som en systematisk feilkilde knyttet til gjenkalling av tidligere eksponeringer, eksempelvis fysisk aktivitet (Coughlin, 1990).

Det er gjennomført et fåtall studier med både varierte metoder og utvalg på effekten av social desirability på selvrapportert fysisk aktivitet. Evidensen på hvorvidt social desirability har en effekt på selvrapportert aktivitet viser at det er variasjon om hvorvidt fenomenet har en effekt eller ikke (Adams et al., 2005; Crutzen & Göritz, 2011; Jago, Thompson, Baranowski, Cullen, & Baranowski, 2006; Motl, McAuley, & DiStefano, 2005). En essensiell bemerkning med disse studiene er at utvalgene er så varierte at det er utfordrende å skulle sammenligne disse, i tillegg

er ulike skjemaer på måling av grad av social desirability benyttet. Studiene kan likevel gi indikasjoner på at det eksiterer en svak relasjon mellom social desirability og selvrapporing av fysisk aktivitet. Crutzen and Göritz (2011) er en av studiene hvor man ikke finner noen relasjon mellom social desirability og fysisk aktivitet. Denne studien bygger på data fra en web-basert undersøkelse med IPAQ-SF. Forfatterne argumenterer med at web-undersøkelsen er med på å skape en sosial distanse mellom undersøker og deltaker. I dette prosjektet var det 10 deltakere som hadde direkte kontakt med student, som er både undersøker og forfatter, og den sosiale distansen mellom undersøker og deltaker er derfor minimal. Det er ikke gjort kartlegging av recall bias og social desirability i dette masterprosjektet, men man kan likevel ikke se bort i fra at social desirability kan ha hatt en innvirkning på selvrapporing av fysisk aktivitet til utvalget i dette prosjektet.

5.2.5 Aktivitetsdata fra måling med Axivity AX3

Som nevnt tidligere i punkt 2.2.5 i teori er det gjennomført en del studier på validiteten til AX3 som tyder på akseptabel validitet. En nylig publisert studie av White et al. (2019) publisert 2. april 2019 viste at korrelasjon mellom doubly labeled water og måling med AX3 på håndledet var 0.67 og 0.91 ved måling av henholdsvis energiforbruk under aktivitet og totalt energiforbruk. Det betyr at AX3, avhengig av måling på energiforbruk under aktivitet og totalt energiforbruk, tyder på moderat til høy validitet. Dette er en svakhet som må tas i betraktning ved tolkning av resultatene på den samtidige validiteten til IPAQ-SF som fremkommer i dette masterprosjektet. Som nevnt i teorien ble det også gjennomført en valideringsstudie av AX3 opp mot et annet tri-aksialt akselerometer ved navn RT3 på 23 eldre med ulik grad av fysiske funksjonsnedsettelse som viste en moderat korrelasjon (0.69) mellom AX3 og RT3 (Clarke et al., 2017). RT3 har imidlertid blitt validert opp mot gullstandarden doubly labeled water blant et utvalg voksne hvor man fant at RT3 underestimerte aktivitetsrelatert energiforbruk med gjennomsnittlig 15% (Maddison et al., 2009). Lite utvalg (n=23) i studien av Clarke et al. (2017) gjør at resultatene må tolkes med varsomhet.

I måling av fysisk aktivitet med akselerometer er det ulike forhold tilknyttet målemetoden som i større eller mindre grad kan påvirke resultatet av målingene. I dette prosjektet var det både under innsamling, prosessering og analysing av dataene grunnlag for feilkilder. I avsnittene nedenfor blir disse potensielle feilkildene belyst og diskutert i lys av gjeldende litteratur.

Måling med akselerometer i seg selv er ikke uproblematisk. Ved måling av fysisk aktivitet ønsker man i størst mulig grad at målingene gjenspeiler de faktiske nivåene av aktivitet til deltakerne. Det er imidlertid evidens som tyder på at man oppnår en reaktivitet i form av økt aktivitet ved bruk av akselerometer (Baumann et al., 2018; C. Cooper et al., 2018; Goode et al., 2017). Baumann et al. (2018) hevder at faktorer som økt motivasjon og social desirability kan være mulige årsaksforklaringer på hvorfor man ser tendenser til reaktivitet ved bruk av akselerometer. Det er usikkerhet knyttet til hvor stor effekt reaktivitet har på akselerometermåling, i tillegg har man ikke lagt til rette for å undersøke reaktivitet i dette prosjektet. Derfor er det usikkert hvorvidt dette kan ha påvirket resultatene i dette masterprosjektet.

En annen potensiell feilkilde ved måling med akselerometer er bruk av kjøretøy. I en studie på validering av ActiGraph på 35 bussjåfører fant de store forskjeller i aktivitetsdata fra akselerometrene på arbeidsdager og fridager, hvor det var høyest aktivitet på arbeidsdagene. Dette kan indikere at akselerasjon fra kjøretøy er opphav til falske positiver i måling av fysisk aktivitet (Varela Mato, Yates, Stensel, Biddle, & Clemes, 2017). Grunnet den sparsomme evidensen som eksisterer på dette området, er det usikker hvor stor påvirkning kjøretøy faktisk har på akselerometermåling. Likevel kan det være viktig å ta dette i betraktning i vurdering av aktivitetsnivå målt med akselerometer. Det er fra tidligere gjennomført en studie på å skape automatiske algoritmer for å oppdage bilkjøring med gode resultater (Strackiewicz, Urbanek, Fadel, Crainiceanu, & Harezlak, 2016). Det ble imidlertid i dette masterprosjektet ikke tatt høyde for påvirkning av bilkjøring siden det ikke er lagt til rette for dette i prosesseringsmetodene som er benyttet. Det er derfor en mulighet for at bilkjøring kan ha påvirket målingen av fysisk aktivitet med akselerometer i dette masterprosjektet.

5.2.6 Metode1Esliger og Metode2Doherty

For å prosessere rådata fra Axivity AX3 ble det i dette masterprosjektet benyttet to ulike prosesseringsmetoder som har sitt grunnlag fra tidligere studier av Esliger et al. (2011), A. Doherty et al. (2017) og (Willettts et al., 2018). Nedenfor følger diskusjon på utfordringer knyttet til disse prosesseringsmetodene. Årsaken til at man ønsket å benytte to ulike prosesseringsmetodene i dette masterprosjektet var nettopp fordi man ønsket å undersøke om dette hadde påvirkning en påvirkning på rådata fra akselerometer i form av endring i antall minutter med fysisk aktivitet. Således viste resultatene at det var en forskjell mellom de to prosesseringsmetodene som var benyttet.

Per i dag foregår det et skifte fra prosessering av akselerometerdata med algoritmer bygd på grenseverdier, over til algoritmer som i stedet for bruker maskinlæring (Farrahi, Niemela, Kangas, Korpelainen, & Jamsa, 2019). Eldre prosesseringsmetoder på akselerometer-data produserer utelukkende data på intensiteten til aktiviteten og ikke type aktivitet. Med maskinlæringsalgoritmer har man mulighet til å kunne si noe om hva slags aktivitet som har blitt målt, eksempelvis gange eller sykling (Willetts et al., 2018).

Den primære forskjellen mellom Metode1_{Esliger} og Metode2_{Doherty} er altså type algoritme som benyttes for å prosessere rådataene fra akselerometer. Metode1_{Esliger} på en mindre komplisert algoritme med beregning av Signal vector magnitude som igjen klassifiseres i ulike intensiteter som igjen er basert på MET-verdier fra studien av Esliger et al. (2011). Utgangspunktet for grenseverdiene som benyttes i Metode1_{Esliger} har sitt opphav i en protokoll hvor bestemte aktiviteter er satt opp i en spesifikk rekkefølge. To studier med utgangspunkt i lignende prosesseringsmetoder viste at en «laboratorie-trent» algoritme ble mer unøyaktig når den ble benyttet i en dagligdags setting og forfatterne av studien mener at akselerometer-data i en mer «dagligdags» kontekst vil variere mer enn det som måles i en laboratorie-basert setting (Gyllensten & Bonomi, 2011). På en annen side er studien av Gyllensten and Bonomi (2011) basert på et utvalg av 20 friske individer, og sammenligning av akselerometer-data ble gjort opp mot intelligent device for energy expenditure and physical activity (IDEEA) og aktivitetsdagbok. Som kriteriestandard er det usikkert hvorvidt bruken av IDEEA og aktivitetsdagbok gir valide data i dagligdags kontekst siden det oppgis at IDEEA er kun validert i laboratorie-studier (Gyllensten & Bonomi, 2011). Indikasjonene på høyere grad av unøyaktighet ved bruk av laboratorie-baserte prosesseringsmetoder i en dagligdags-setting er likevel tilstede og gir en pekepinn på nøyaktigheten til Metode1_{Esliger}.

Det finnes mange ulike merker og modeller av akselerometre på markedet som tenkes å gi en del utfordringer knyttet til måling med akselerometer (Wijndaele et al., 2015). For Metode1_{Esliger} er det en mulig feilkilde at man har benyttet AX3 som akselerometer, da metoden bygger på studier gjennomført med GENEAA akselerometer. På den andre siden er det evidens for at GENEAA og AX3 er å regne som ekvivalent i estimering av fysisk aktivitet (Rowlands et al., 2018). Det er derfor grunnlag for å tro at eventuelle feilkilder ved at AX3 ble brukt er små.

Et annet aspekt som kan være en potensiell feilkilde i Metode1_{Esliger} er målehastigheten målt i hertz (Hz). I studien av Esliger et al. (2011) ble det benyttet 80 Hz målehastighet på GENEAA-

akselerometrene. Siden man med AX3-akselerometrene hadde valget mellom 25, 50, 100, 200 eller 400, falt valget på 100 Hz, som ifølge produsenten kunne gi en batterikapasitet på 2 uker (www.axivity.com). Ifølge Migueles et al. (2017) er anbefalt målehastighet for voksne og eldre voksne angitt som 90-100 Hz. Det finnes evidens som gir antydning til å tro at ulik målehastighet kan ha en påvirkning på akselerometerdata, hvor man så at en økning i antall minutter aktivitet med høyere målehastighet (Brond & Arvidsson, 2016). Det er utfordrende å skulle fastslå hvor står påvirkning målehastighet på 80 til 100 Hz gir på antall minutter aktivitet, men det er rimelig å anta at det er snakk om for liten forskjell til at det vil kunne påvirke dataene på aktivitet fra akselerometer.

I form av epoker var det i Metode1_{Esliger} kun mulighet for en oppløsning på 60-sekunder epoker i motsetning til 30 sekunders epoker ved benyttelse av Metode2_{Doherty} (A. Doherty et al., 2017). Her er det ifølge Migueles et al. (2017) ingen studier som rapporterer om innflytelsen ulike epoke-tider har på aktivitetsmåling hos voksne og voksne eldre. Det betyr at man i hovedsak ikke har en standardisering av hvilken epoke-tid som er optimalt for klassifisering av aktivitet hos voksne og voksne eldre. Imidlertid viste en studie på 401 tenåringer at en endring fra 3 til 60 sekunders epoker førte til en gradvis nedgang i antall minutter moderat-til-anstrengende aktivitet (Aibar et al., 2014). Til tross for at det er rimelig å anta at tenåringer har et annet aktivitetsmønster enn det eldre har, er det ikke utenkelig at man kan finne lignende resultater hos eldre utvalg. Som tidligere nevnt er Metode1_{Esliger} basert på en laboratorie-studie hvor man har spesifikke protokoller over hvilke aktiviteter som skal utføres og hvor lenge. Det vil si at man innenfor disse kontrollerte settingene kan gå ut ifra at innenfor 1 minuts-epoken er aktiviteten rimelig lik. Dette kan tenkes å være utfordrende når man skal overføre dette til mer ukontrollerte settinger ved måling av dagligdags fysisk aktivitet hvor man i 10-20 sekunder gjør husarbeid, mens man i det neste øyeblikk er stillesittende. Her trengs det mer forskning både for å kunne fastsette standarder på akselerometer-måling av fysisk aktivitet til eldre, men også for å kunne ta høyde for eventuelle feilkilder som eksisterer i akselerometer-måling.

I vurderingen av metodisk kvalitet knyttet til Metode2_{Doherty} er det noen forskjeller i forhold til vurderingen av Metode1. Ulike metodiske valg som er knyttet til bruk av samme type akselerometer, lik plassering av akselerometer og samme målehastighet og sensitivitet gjør at man har forholdt seg i større grad til metoden som er gjennomført i studiene av A. Doherty et al. (2017) og Willetts et al. (2018). Eventuelle feilkilder som skyldes ulikheter i metode er derfor blitt minimerte. Det betyr likevel ikke at det ikke er feilkilder i Metode2.

Valideringen av prosesseringsmetoden, som er grunnlaget for Metode2^{Doherty}, har blitt gjennomført med bærbart kamera. Hvorvidt bærbart kamera er en adekvat kriteriestandard er noe som kan diskuteres. Til tross for at Metode2^{Doherty} baserer seg på metoden til en studie med over 90 000 deltakere, er denne ikke validert opp mot gullstandarden doubly labeled water (A. Doherty et al., 2017; Willetts et al., 2018). I studien av Willetts et al. (2018) blir nøyaktigheten til prosesseringsmetoden sammenlignet med bærbart kamera festet på høyde med øynene til 132 deltakere. Disse bildene som var datostemplet ble så sammenlignet med data fra akselerometer og deretter klassifisert ut ifra ulike intensiteter av fysisk aktivitet. En studie på lignende type bærbart kamera, sammenlignet med akselerometer viste at 19 % av bildene fra det bærbare kameraet var ikke egnet til klassifisering av aktivitet (A. R. Doherty et al., 2013). En annen svakhet med kameraene som blir benyttet både i Willetts et al. (2018) og A. R. Doherty et al. (2013) er at de bilder omtrent hvert 20 sekund. Her kan man tenke seg til at eksempelvis sportslige aktivitet hvor intensitetsnivået kan variere i stor grad innenfor en tidsramme på 20 ikke vil kunne la seg klassifisere på adekvat måte. Som kriteriestandard er det derfor usikkert om denne kan være en god nok metode for å kunne validere prosesseringsmetoder på akselerometer.

Spørsmålene i IPAQ-SF på samtlige intensiteter (gange, moderat og meget anstrengende) tar forbehold om at respondenten kun oppgir aktivitet som har vart 10 minutter eller mer (punkt 2.2.3). Dette gjorde at man, for å ha et likt sammenligningsgrunnlag mellom aktivitetsdata fra akselerometer og IPAQ-SF, anså det som hensiktsmessig å hente ut intervaller på 10 minutter eller mer på aktivitetsdata fra prosesseringsmetodene. Inndeling i 10-minutters intervaller ble også gjennomført i valideringsstudien av Grimm et al. (2012). Siden det ikke var mulig å hente ut aktivitetsdata på 10 minutter eller automatisk via egnet programvare, måtte dette gjøres manuelt av masterstudenten. Den manuelle behandlingen av aktivitetsdata i etterkant av prosessering av akselerometerdata kan derfor ha ført til unøyaktigheter i aktivitetsdataene fra akselerometer. Gitt den store mengden med data som ble samlet inn på hver enkelt deltaker, kan det ikke utelukkes at noen data gikk enten tapt eller at det skjedde en feilberegning. Som kontroll av mulig unøyaktigheter i denne delen av behandling av akselerometer-data ble derfor de gjennomførte beregningene gjennomgått to ganger og samme resultat fremkom begge ganger.

5.2.7 Feasibility

For vurdering av deltakernes erfaringer ved måling med både IPAQ-SF og AX3 akselerometer ble et spørreskjema adaptert fra tidligere gjennomførte studier på yngre utvalg. Det at man ikke har benyttet et standardisert spørreskjema i dette prosjektet gjør at resultatene fra spørreskjemaet på feasibility blir betydelig svekket. Det ideelle hadde vært om man hadde funnet et skjema som dekket alle behov, eller gjennomført en utprøving av skjemaet i forkant av selve datainnsamlingen. En annen utfordring knyttet til bruk av dette spørreskjemaet er at spørsmålene kan ha blitt mistolket av deltakerne. Spørreskjemaet gir likevel noen indikasjoner på hva deltakerne har erfart rundt bruken av IPAQ-SF og AX3 akselerometer.

Et annet punkt det er viktig å belyse i innsamlingen av erfaringene til deltakerne er at tidspunktet for når de svarte på spørsmålene varierte mellom deltakerne. Det betyr at noen av deltakerne (n=5) fikk spørsmålene samtidig med IPAQ-SF-skjemaet, noe som gjorde at de kunne svare på spørsmålene direkte etter at måleperioden med akselerometer var ferdig. I motsetning til for deltakerne som først ble rekruttert til masterprosjektet (n=5) ble disse spørsmålene hentet inn noen måneder etter da man så, grunnet svært treg rekruttering, at man måtte gjøre en justering på masterprosjektet. Dette kan tenkes at den lange perioden mellom måling med IPAQ-SF og AX3 akselerometer og besvarelse av spørreskjemaet kan ha bidratt til bias i besvarelsen fra de fem deltakerne som fikk spørsmålene i etterkant.

5.3 Kliniske implikasjoner / videre forskning

Med grunnlag i vurderingen som er gjort på validering og feasibility er det hensiktsmessig å vurdere hvilke kliniske implikasjoner resultatene i dette prosjektet har. For å kunne diskutere implikasjoner er klinisk erfaring, tidligere evidens og resultatene fra dette prosjektet satt som rammeverk for hvordan disse kan implementeres i praksis.

Både dette masterprosjektet og andre valideringsstudier tyder på at IPAQ-SF gjennomgående har lav grad av validitet, til tross for at det ved en måling ble funnet tilfredsstillende grad av validitet. For kartlegging av fysisk aktivitetsnivå er objektive målemetoder som akselerometer et alternativ som kan tenkes å ha en plass i den kliniske hverdagen. Fordelene med akselerometer er dessuten at man unngår uten potensiell bias i form av social desirability og recall bias. Derimot er et av ulempene ved bruk av akselerometer at man ikke kan si noe om konteksten til aktiviteten som blir målt (Sylvia et al., 2014). Eksempelvis for yrkesaktive er det

evidens som tyder på at meget anstrengende fysisk aktivitet på jobb ikke er gunstig for muskel- og skjeletthelsen (I. Heuch, Heuch, Hagen, & Zwart, 2017; Søgaard & Sjøgaard, 2017). I en fysioterapeutisk praksis vil det derfor kunne være viktig at man kan skille mellom fysisk aktivitet på fritiden og jobb for å kunne hjelpe til med å regulere aktivitetsnivået på fritid og arbeidet til pasienten. I lys av resultatene fra dette prosjektet så man at Metode2^{Doherty} var en komplisert prosesseringsmetode å gjennomføre. Det blir derfor viktig at man finner metoder som ikke bare har gode måleegenskaper, men at de også er enkle i bruk.

Dersom helsepersonell i møte med pasienter med ryggsmarter skal gi anbefalinger om fysisk aktivitet trenger vi noen gode metoder for å kunne måle hvorvidt pasientene har fulgt opp anbefalingen som er gitt, og hvorvidt det har vært en økning av betydning. utfordringen ligger i standardiseringen av hvordan vi måler aktivitet med akselerometer, og hvordan vi prosesserer disse dataene. En pågående diskusjon er behovet for standardisering av metoder for å måle fysisk aktivitet (Steene-Johannessen et al., 2016; Richard P. Troiano, 2009). Utviklingen av nye metoder og kriterier innen måling av fysisk aktivitet, og ikke minst ulik rapportering av hvordan måling med akselerometer, har bidratt til at det er utfordrende å enes om en standard og det argumenteres at det er en generelt lav generaliserbarhet mellom studiene som benytter akselerometer (Miguelles et al., 2017; Montoye et al., 2018). Det er derfor nødvendig med videre forskning på prosesseringsmetoder på rådata fra akselerometer og det er hensiktsmessig å utarbeide en standard protokoll for måling av fysisk aktivitet med akselerometer.

6.0 Konklusjon

Problemstilling og forskningsspørsmålene for dette masterprosjektet var: «*Er det samsvar mellom selvrapportert fysisk aktivitet på IPAQ-SF, og objektiv målt fysisk aktivitet ved bruk av AX3 hos eldre pasienter med ryggsmarter?*»

Prosjektet ønsket derfor å undersøke følgende forskningsspørsmål:

Hvordan er den samtidige validiteten til IPAQ-SF hos pasienter >55 år med ryggsmarter?

Hvordan er feasibility til henholdsvis IPAQ-SF og AX3 akselerometer?

I tillegg ble følgende spørsmål eksplorert:

- Møter utvalget som er undersøkt kravene for fysisk aktivitet (150 minutter moderat fysisk aktivitet/75 minutter anstrengende fysisk aktivitet)?
- Har valg av prosesseringsmetode på akselerometerdata innvirkning på antall minutter fysisk aktivitet hos deltakerne?

Samsvaret mellom selvrapportert fysisk aktivitet på IPAQ-SF og objektiv målt fysisk aktivitet med bruk av AX3 akselerometer viser en høy korrelasjon mellom IPAQ-SF og AX3 akselerometer ved måling av moderat fysisk aktivitet og henholdsvis lav til ingen korrelasjon ved måling av gange og meget anstrengende fysisk aktivitet. På grunnlag av disse resultatene gir det grunn til å forkaste nullhypotesen som ble fremsatt i dette masterprosjektet. Videre ble både IPAQ-SF og AX3 akselerometer av deltakerne i dette prosjektet ansett som akseptable målemetoder på fysisk aktivitet. Imidlertid var det en del tapte data fra IPAQ-SF, sammenlignet med AX3 som ga høyt antall valide måledøgn, som gir grunn til å foretrekke AX3. For de eksplorerende spørsmålene viste data at 70% av deltakerne ved selvrapportering via IPAQ-SF oppnådde kravene om 150 minutter moderat fysisk aktivitet i uka, mens 100 og 90 % tilfredsstilte kravene når måling ble gjort med henholdsvis Metode1^{Esliger} og Metode2^{Doherty}. Korrelasjonsanalyser viste at det var diskrepans mellom Metode1^{Esliger} og Metode2^{Doherty} ved prosessering av rådata fra akselerometer, noe som kan tyde på en forskjell i hvordan disse prosesseringsmetodene klassifiserer fysisk aktivitet målt i antall minutter. Kombinert med tapte data fra IPAQ-SF og tidligere evidens på validiteten til IPAQ-SF gir dette grunnlag for å ikke anbefale IPAQ-SF videre som metode på måling av fysisk aktivitet for dette utvalget. Lavt antall deltakere gjør at resultatene må tolkes med forsiktighet.

7.0 Appendiks

Vedlegg 1.

Fysisk aktivitet

Vi er interessert i informasjon om ulike former for fysisk aktivitet som folk driver med i dagliglivet.

Spørsmålene gjelder tiden du har brukt på fysisk aktivitet de siste 7 dagene. Vennligst svar på alle spørsmålene uansett hvor fysisk aktiv du selv synes du er. Tenk på aktiviteter du gjør på jobb, som en del av hus- og hagearbeid, for å komme deg fra et sted til et annet, og aktiviteter på fritiden (rekreasjon, mosjon og sport).

Tenk på all meget anstrengende aktivitet du har drevet med de siste 7 dagene. Meget anstrengende aktivitet som krever hard innsats og får deg til å puste mye mer enn vanlig. Ta bare med aktiviteter som varer minst 10 minutter i strekk.

207. Hvor mange dager i løpet av de siste 7 dagene har du drevet med meget anstrengende fysisk aktivitet som tunge løft, gravearbeid, aerobics løp eller rask sykling?

1. _____ Dager

2. Ingen meget anstrengende aktivitet (gå til spørsmål 209)

208. Hvor lang tid bruker du vanligvis på meget anstrengende fysisk aktivitet på en av disse dagene?

_____ Timer per dag _____ Minutter per dag _____ Vet ikke/usikker

Tenk på all middels anstrengende aktivitet du har drevet med de siste 7 dagene. Middels anstrengende aktivitet er aktivitet som krever moderat innsats og får deg til å puste litt mer enn vanlig. Ta bare med aktiviteter som varer minst 10 minutter i strekk.

209. Hvor mange dager i løpet av de siste 7 dagene har du drevet med middels anstrengende fysisk aktivitet som å bære lette ting, jogge eller sykle i moderat tempo? Ikke ta med gange.

1. _____ Dager

2. Ingen middels anstrengende aktivitet (gå til spørsmål 211)

210. Hvor lang tid brukte du vanligvis på middels anstrengende fysisk aktivitet på en av disse dagene?

_____ Timer per dag _____ Minutter per dag _____ Vet ikke/usikker

Tenk på tiden du har brukt på å gå de siste 7 dagene. Dette inkluderer gange på jobb og hjemme, gange fra sted til et annet eller gange som du gjør på tur eller som trening på fritiden.

211. Hvor mange dager i løpet av de siste 7 dagene gikk du i minst 10 minutter i strekk?

1. _____ Dager

2. _____ Gikk ikke (gå til spørsmål 213)

212. Hvor lang tid brukte du vanligvis på å gå på en av disse dagene?

_____ Timer per dag _____ Minutter per dag _____ Vet ikke/usikker

Det neste spørsmålet omfatter all tid du tilbrakte sittende på ukedagene i løpet av de siste 7 dagene. Inkluder tid du har brukt på å sitte på jobb, hjemme, på kurs og på fritiden. Dette kan tilsvare tiden du sitter ved et arbeidsbord, hos venner, mens du leser, eller sitter eller ligger for å se på TV.

213. Hvor lang tid brukte du på å sitte på en vanlig hverdag i løpet av de siste 7 dagene?

_____ Timer per dag _____ Minutter per dag _____ Vet ikke/usikker

Vedlegg 2

Nedenfor er en rekke uttalelser knyttet til bruk av aktivitetsarmbåndet og spørreskjemaet som ble fylt ut. Sett strek under ett av svaralternativene på skalaen fra 1-5.

1 «Jeg hadde på meg aktivitetsarmbåndet hele tiden (7 dager) 1. Helt enig 2. Enig 3. Nøytral 4. Uenig 5. Helt uenig
2 «Jeg ble mer aktiv av å bære aktivitetsarmbåndet» 1. Helt enig 2. Enig 3. Nøytral 4. Uenig 5. Helt uenig
3 «Aktivitetsarmbåndet hindret meg i å gjøre hverdagslige gjøremål» 1. Helt enig 2. Enig 3. Nøytral 4. Uenig 5. Helt uenig
4 «Det var ubehagelig å ha på seg aktivitetsarmbåndet» 1. Helt enig 2. Enig 3. Nøytral 4. Uenig 5. Helt uenig
5 «Jeg kunne gått med aktivitetsarmbåndet på nytt igjen» 1. Helt enig 2. Enig 3. Nøytral 4. Uenig 5. Helt uenig

6 «Jeg besvarte ut hele IPAQ-skjemaet» 1. Helt enig 2. Enig 3. Nøytral 4. Uenig 5. Helt uenig
7 «Jeg ble mer aktiv av å skulle besvare IPAQ-skjemaet» 1. Helt enig 2. Enig 3. Nøytral 4. Uenig 5. Helt uenig
8 «Det var lett å forstå spørsmålene i IPAQ-skjemaet» 1. Helt enig 2. Enig 3. Nøytral 4. Uenig 5. Helt uenig
9 «Det var vanskelig å tilbakekalle siste 7 dagers fysiske aktivitet» 1. Helt enig 2. Enig 3. Nøytral 4. Uenig 5. Helt uenig
10 «Jeg kunne besvart IPAQ-skjemaet på nytt igjen» 1. Helt enig 2. Enig 3. Nøytral 4. Uenig 5. Helt uenig

Vedlegg 3.

Vil du delta i forskningsprosjektet

” Samsvar mellom selvrapportert og objektiv fysisk aktivitet hos eldre med ryggplager – en del av BACE-N”?

Dette er et spørsmål til deg om å delta i et forskningsprosjekt hvor formålet er å undersøke sammenhengen mellom selvrapportert fysisk aktivitet målt med spørreskjemaet IPAQ (International Physical Activity Questionnaire Short Form) og objektivt målt gjennom bruk av AX3 akselerometer (aktivitetsmåler). Prosjektet er en del av et internasjonalt forskningsprosjekt som kartlegger hvordan det går med personer > 55 år som opplever en ryggepisode (Back Pain in Elderly; BACE-prosjektet). I dette skrivet gir vi deg informasjon om målene for prosjektet og hva deltakelse vil innebære for deg.

Undertegnede er fysioterapeut og masterstudent ved OsloMet og skal gjennom dette masterprosjekt sammenligne svarene fra spørreskjemaet IPAQ med målinger av fysisk aktivitetsnivå ved bruk av aktivitetsmålere som deltakerne bærer på håndleddet i 7 dager.

Formål

Formålet med studien er å undersøke om IPAQ-skjemaet er et egnet spørreskjema til å fange opp fysisk aktivitet.

Hvem er ansvarlig for forskningsprosjektet?

Student: Emil Thomassen – telefon: 95 25 22 16, E-post: s326902@oslomet.no

Veileder: Kjersti Storheim – telefon: 99 70 87 83, E-post: kjersti.storheim@medisin.no

Hvorfor får du spørsmål om å delta?

Du forespørres om deltakelse fordi du er inkludert i BACE-N-studien. I BACE-N-studien svarer du blant annet på et spørreskjema (IPAQ) om ditt fysiske aktivitetsnivå. Man vet lite om aktivitetsnivået til voksne eldre med ryggmerter.

Kontaktinfo er hentet fra stipendiat Rikke Munk Killingmo som videresender godkjenninger fra deltakere til å delta i prosjektet.

Hva innebærer det for deg å delta?

Prosjektet innebærer at du fyller ut spørreskjemaet IPAQ to ganger og at du går med en aktivitetsmåler på håndleddet i 7 dager. De som samtykker til deltakelse i masterprosjektet møter med masterstudent i forbindelse med et tidspunkt de likevel svarer på spørreskjema i BACE-N. Deltakere utstyres med Axivity AX3 akselerometer og får utdelt et ekstra spørreskjema. Aktivitetsmåleren er på størrelsen med en klokke og festet på håndleddet med en klokkeim av silikon og skal bæres hele døgnet i 7 dager. Denne tåler vann inntil 1.5 meter. Deltakere til masterprosjekt rekrutteres enten gjennom førstegangsmåling, 3, 6, 12 eller 24 måneder ut i BACE-N-prosjektet.

Tidsaspektet blir derfor seende slik ut: Ved samtykke til deltakelse blir du kontaktet av masterstudent hvor det lages en avtale enten at du kommer til OsloMET for montering av aktivitetsmåler og utfylling av BACE-N-skjemaet, eller at masterstudent kan komme hjem til deg. Du fyller deretter ut skjemaet som hører til BACE-N. Deltakere går med

aktivitetsarmbåndet i 7 døgn. Når armbåndet tas av etter 7 døgn fylles nytt IPAQ-skjema ut og begge deler (armbånd og IPAQ-spørreskjema) returneres til masterstudent enten ved personlig oppmøte, eller ved å returnere det i en forhåndsfrankert returkonvolutt.

I prosjektet vil vi innhente og registrere opplysninger om deg. Disse opplysningene går på hvor fysisk aktiv du er i løpet gjennom en vanlig uke i form av din egen rapportering i IPAQ-skjemaet. I dette prosjektet er vi ute etter å se hvor nøyaktig IPAQ-skjemaet er til å måle fysisk aktivitet, ikke hvor fysisk aktiv du er. Derfor er det viktig at du prøver å opprettholde ditt normale aktivitetsnivå som vanlig.

Mulige fordeler og ulemper med å delta:

- Ulemper med prosjektet er at det kan være knyttet noe ubehag til å ha festet aktivitetsmåleren til håndleddet. Denne aktivitetsmåleren kan dusjes med.
- Fordelen med å være med i dette prosjektet er at du kanskje blir bevisst på hvor mye fysisk aktiv du er i løpet av en uke.

Det er frivillig å delta

Det er frivillig å delta i prosjektet. Dersom du ønsker å delta, undertegner du samtykkeerklæringen på siste side. Du kan når som helst og uten å oppgi noen grunn trekke ditt samtykke. Dersom du trekker deg fra prosjektet, kan du kreve å få slettet innsamlede prøver og opplysninger, med mindre opplysningene allerede er inngått i analyser eller brukt i vitenskapelige publikasjoner. Dersom du senere ønsker å trekke deg eller har spørsmål til prosjektet, kan du kontakte Emil Eirik Kvernberg Thomassen på telefon: 95 25 22 16, eller epost: s326902@oslomet.no. Eventuelt kan du kontakt veileder Kjersti Storheim på telefon: 997 08 783 eller epost: kjersti.storheim@medisin.uio.no

Ditt personvern – hvordan vi oppbevarer og bruker dine opplysninger

Vi vil bare bruke opplysningene om deg til formålene vi har fortalt om i dette skrivet. Vi behandler opplysningene konfidensielt og i samsvar med personvernregelverket.

Deltakerne vil ikke kunne gjenkjennes dersom studien publiseres.

Hva skjer med opplysningene dine når vi avslutter forskningsprosjektet?

Prosjektet skal etter planen avsluttes samtidig med BACE-N 25. mars 2025. Dataene blir anonymiserte etter prosjektets slutt, det vil si at personopplysninger slettes. Informasjonen som registreres om deg skal kun brukes slik som beskrevet i hensikten med studien. Du har rett til innsyn i hvilke opplysninger som er registrert om deg og rett til å få korrigeret eventuelle feil i de opplysningene som er registrert.

Alle opplysningene vil bli behandlet uten navn og fødselsnummer eller andre direkte gjenkjennende opplysninger. En kode knytter deg til dine opplysninger gjennom en navneliste.

Prosjektleder har ansvar for den daglige driften av forskningsprosjektet og at opplysninger om deg blir behandlet på en sikker måte. Informasjon om deg vil bli anonymisert eller slettet senest fem år etter prosjektslutt.

Dine rettigheter

Så lenge du kan identifiseres i datamaterialet, har du rett til:

- innsyn i hvilke personopplysninger som er registrert om deg,
- å få rettet personopplysninger om deg,

- få slettet personopplysninger om deg,
- få utlevert en kopi av dine personopplysninger (dataportabilitet), og
- å sende klage til personvernombudet eller Datatilsynet om behandlingen av dine personopplysninger.

Hva gir oss rett til å behandle personopplysninger om deg?

Vi behandler opplysninger om deg basert på ditt samtykke.

På oppdrag fra OsloMet har NSD – Norsk senter for forskningsdata AS vurdert at behandlingen av personopplysninger i dette prosjektet er i samsvar med personvernregelverket.

Prosjektet er vurdert av Regional komite for medisinsk og helsefaglig forskningsetikk Øst, men er klassifisert som en kvalitetsstudie.

Hvor kan jeg finne ut mer?

Hvis du har spørsmål til studien, eller ønsker å benytte deg av dine rettigheter, ta kontakt med:

- *OsloMet* ved Margreth Grotle (tlf: +4767236043), Kjersti Storheim (997 08 783), eller Emil Thomassen (tlf: 95 25 22 16)
- Vårt personvernombud: NSD – Norsk senter for forskningsdata AS, på epost (personverntjenester@nsd.no) eller telefon: 55 58 21 17.

Med vennlig hilsen

Kjersti Storheim
(Forsker/veileder)

Emil Thomassen
Student

Samtykkeerklæring

Jeg har mottatt og forstått informasjon om prosjektet Samsvar mellom selvrapportert og objektiv fysisk aktivitet hos eldre med ryggplager – en del av BACE-N, og har fått anledning til å stille spørsmål. Jeg samtykker til:

- å delta i måling med aktivitetsarmbånd
- å delta i utfylling av IPAQ-spørreskjema
- å få lagret opplysninger om meg frem til prosjektets slutt

Jeg samtykker til at mine opplysninger behandles frem til prosjektet er avsluttet, 1. mars 2025.

(Signert av prosjektdeltaker, dato)

8.0 Referanser

- Adams, S. A., Matthews, C. E., Ebbeling, C. B., Moore, C. G., Cunningham, J. E., Fulton, J., & Hebert, J. R. (2005). The effect of social desirability and social approval on self-reports of physical activity. *Am J Epidemiol*, *161*(4), 389-398. doi:10.1093/aje/kwi054
- Aibar, A., Bois, J. E., Zaragoza, J., Generelo, E., Julian, J. A., & Paillard, T. (2014). Do epoch lengths affect adolescents' compliance with physical activity guidelines? *J Sports Med Phys Fitness*, *54*(3), 326-334.
- Ainsworth, B. E., Haskell, W. L., Herrmann, S. D., Meckes, N., Bassett, D. R., Jr., Tudor-Locke, C., . . . Leon, A. S. (2011). 2011 Compendium of Physical Activities: a second update of codes and MET values. *Med Sci Sports Exerc*, *43*(8), 1575-1581. doi:10.1249/MSS.0b013e31821ece12
- Ainsworth, B. E., Haskell, W. L., Whitt, M. C., Irwin, M. L., Swartz, A. M., Strath, S. J., . . . Leon, A. S. (2000). Compendium of physical activities: an update of activity codes and MET intensities. *Med Sci Sports Exerc*, *32*(9 Suppl), S498-504.
- Allan, D. B., & Waddell, G. (1989). An historical perspective on low back pain and disability. *Acta Orthopaedica Scandinavica*, *60*(sup234), 1-23. doi:10.3109/17453678909153916
- Bahr, R., & Helsedirektoratet (Eds.). (2009). *Aktivitetshåndboken : fysisk aktivitet i forebygging og behandling*. Oslo: Helsedirektoratet.
- Bassett, D. R., Troiano, R. P., McClain, J. J., & Wolff, D. L. (2015). Accelerometer-based physical activity: total volume per day and standardized measures. *Med Sci Sports Exerc*, *47*(4), 833-838. doi:10.1249/mss.0000000000000468
- Bauman, A., Bull, F., Chey, T., Craig, C. L., Ainsworth, B. E., Sallis, J. F., . . . Pratt, M. (2009). The International Prevalence Study on Physical Activity: results from 20 countries. *Int J Behav Nutr Phys Act*, *6*, 21. doi:10.1186/1479-5868-6-21
- Baumann, S., Gross, S., Voigt, L., Ullrich, A., Weymar, F., Schwaneberg, T., . . . Ulbricht, S. (2018). Pitfalls in accelerometer-based measurement of physical activity: The presence of reactivity in an adult population. *Scand J Med Sci Sports*, *28*(3), 1056-1063. doi:10.1111/sms.12977
- Bernstein, I. A., Malik, Q., Carville, S., & Ward, S. (2017). Low back pain and sciatica: summary of NICE guidance. *BMJ*, *356*, i6748. doi:10.1136/bmj.i6748
- Boonstra, A. M., Stewart, R. E., Köke, A. J. A., Oosterwijk, R. F. A., Swaan, J. L., Schreurs, K. M. G., & Schiphorst Preuper, H. R. (2016). Cut-Off Points for Mild, Moderate, and

Severe Pain on the Numeric Rating Scale for Pain in Patients with Chronic Musculoskeletal Pain: Variability and Influence of Sex and Catastrophizing. *Frontiers in psychology*, 7, 1466-1466. doi:10.3389/fpsyg.2016.01466

- Bowen, D. J., Kreuter, M., Spring, B., Cofta-Woerpel, L., Linnan, L., Weiner, D., . . . Fernandez, M. (2009). How we design feasibility studies. *American journal of preventive medicine*, 36(5), 452-457. doi:10.1016/j.amepre.2009.02.002
- Brage, S., Ihlebaek, C., Natvig, B., & Bruusgaard, D. (2010). [Musculoskeletal disorders as causes of sick leave and disability benefits]. *Tidsskr Nor Laegeforen*, 130(23), 2369-2370. doi:10.4045/tidsskr.10.0236
- Bressler, H. B., Keyes, W. J., Rochon, P. A., & Badley, E. (1999). The prevalence of low back pain in the elderly. A systematic review of the literature. *Spine (Phila Pa 1976)*, 24(17), 1813-1819.
- Brond, J. C., & Arvidsson, D. (2016). Sampling frequency affects the processing of Actigraph raw acceleration data to activity counts. *J Appl Physiol (1985)*, 120(3), 362-369. doi:10.1152/jappphysiol.00628.2015
- Buchbinder, R., van Tulder, M., Oberg, B., Costa, L. M., Woolf, A., Schoene, M., & Croft, P. (2018). Low back pain: a call for action. *Lancet*, 391(10137), 2384-2388. doi:10.1016/s0140-6736(18)30488-4
- Carter, R. E., & Lubinsky, J. (2016). *Rehabilitation research : principles and applications*. In: Carvalho, F. A., Maher, C. G., Franco, M. R., Morelhaio, P. K., Oliveira, C. B., Silva, F. G., & Pinto, R. Z. (2017). Fear of Movement Is Not Associated With Objective and Subjective Physical Activity Levels in Chronic Nonspecific Low Back Pain. *Arch Phys Med Rehabil*, 98(1), 96-104. doi:10.1016/j.apmr.2016.09.115
- Caspersen, C. J., Powell, K. E., & Christenson, G. M. (1985). Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. *Public Health Reports*, 100(2), 126-131.
- Chiarotto, A., Boers, M., Deyo, R. A., Buchbinder, R., Corbin, T. P., Costa, L. O. P., . . . Ostelo, R. W. (2018). Core outcome measurement instruments for clinical trials in nonspecific low back pain. *Pain*, 159(3), 481-495. doi:10.1097/j.pain.0000000000001117
- Clarke, C. L., Taylor, J., Crighton, L. J., Goodbrand, J. A., McMurdo, M. E. T., & Witham, M. D. (2017). Validation of the AX3 triaxial accelerometer in older functionally impaired people. *Aging Clin Exp Res*, 29(3), 451-457. doi:10.1007/s40520-016-0604-8

- Cleland, C., Ferguson, S., Ellis, G., & Hunter, R. F. (2018). Validity of the International Physical Activity Questionnaire (IPAQ) for assessing moderate-to-vigorous physical activity and sedentary behaviour of older adults in the United Kingdom. *BMC Med Res Methodol*, *18*(1), 176. doi:10.1186/s12874-018-0642-3
- Cleland, I., Kikhia, B., Nugent, C., Boytsov, A., Hallberg, J., Synnes, K., . . . Finlay, D. (2013). Optimal placement of accelerometers for the detection of everyday activities. *Sensors (Basel)*, *13*(7), 9183-9200. doi:10.3390/s130709183
- Cooper, C., Gross, A., Brinkman, C., Pope, R., Allen, K., Hastings, S., . . . Goode, A. P. (2018). The impact of wearable motion sensing technology on physical activity in older adults. *Exp Gerontol*, *112*, 9-19. doi:10.1016/j.exger.2018.08.002
- Cooper, K., Sani, S., Corrigan, L., MacDonald, H., Prentice, C., Vareta, R., . . . Wiratunga, N. (2017). Accuracy of physical activity recognition from a wrist-worn sensor. *Physiotherapy*, *103*, e47. doi:<https://doi.org/10.1016/j.physio.2017.11.213>
- Copeland, J. L., & Eslinger, D. W. (2009). Accelerometer assessment of physical activity in active, healthy older adults. *J Aging Phys Act*, *17*(1), 17-30.
- Coughlin, S. S. (1990). Recall bias in epidemiologic studies. *J Clin Epidemiol*, *43*(1), 87-91.
- Craig, C. L., Marshall, A. L., Sjostrom, M., Bauman, A. E., Booth, M. L., Ainsworth, B. E., . . . Oja, P. (2003). International physical activity questionnaire: 12-country reliability and validity. *Med Sci Sports Exerc*, *35*(8), 1381-1395. doi:10.1249/01.Mss.0000078924.61453.Fb
- Crutzen, R., & Göritz, A. S. (2011). Does social desirability compromise self-reports of physical activity in web-based research? *Int J Behav Nutr Phys Act*, *8*, 31. doi:10.1186/1479-5868-8-31
- de Vet, H. C. W., Terwee, C. B., Mokkink, L. B., & Knol, D. L. (2011). *Measurement in Medicine: A Practical Guide*. Cambridge: Cambridge University Press.
- Deyo, R. A., Mirza, S. K., & Martin, B. I. (2006). Back pain prevalence and visit rates: estimates from U.S. national surveys, 2002. *Spine (Phila Pa 1976)*, *31*(23), 2724-2727. doi:10.1097/01.brs.0000244618.06877.cd
- Dionne, C. E., Dunn, K. M., Croft, P. R., Nachemson, A. L., Buchbinder, R., Walker, B. F., . . . Von Korf, M. (2008). A consensus approach toward the standardization of back pain definitions for use in prevalence studies. *Spine (Phila Pa 1976)*, *33*(1), 95-103. doi:10.1097/BRS.0b013e31815e7f94
- Doherty, A., Jackson, D., Hammerla, N., Plotz, T., Olivier, P., Granat, M. H., . . . Wareham, N. J. (2017). Large Scale Population Assessment of Physical Activity Using Wrist

- Worn Accelerometers: The UK Biobank Study. *PLoS One*, 12(2), e0169649.
doi:10.1371/journal.pone.0169649
- Doherty, A. R., Kelly, P., Kerr, J., Marshall, S., Oliver, M., Badland, H., . . . Foster, C. (2013). Using wearable cameras to categorise type and context of accelerometer-identified episodes of physical activity. *Int J Behav Nutr Phys Act*, 10, 22.
doi:10.1186/1479-5868-10-22
- Downie, W. W., Leatham, P. A., Rhind, V. M., Wright, V., Branco, J. A., & Anderson, J. A. (1978). Studies with pain rating scales. *Annals of the Rheumatic Diseases*, 37(4), 378.
doi:10.1136/ard.37.4.378
- Ellis, K., Kerr, J., Godbole, S., Lanckriet, G., Wing, D., & Marshall, S. (2014). A random forest classifier for the prediction of energy expenditure and type of physical activity from wrist and hip accelerometers. *Physiol Meas*, 35(11), 2191-2203.
doi:10.1088/0967-3334/35/11/2191
- Engel, G. L. (1977). The need for a new medical model: a challenge for biomedicine. *Science*, 196(4286), 129-136.
- Esliger, D. W., Rowlands, A. V., Hurst, T. L., Catt, M., Murray, P., & Eston, R. G. (2011). Validation of the GENEActiv Accelerometer. *Med Sci Sports Exerc*, 43(6), 1085-1093.
doi:10.1249/MSS.0b013e31820513be
- Falck, R. S., McDonald, S. M., Beets, M. W., Brazendale, K., & Liu-Ambrose, T. (2016). Measurement of physical activity in older adult interventions: a systematic review. *Br J Sports Med*, 50(8), 464-470. doi:10.1136/bjsports-2014-094413
- Farrahi, V., Niemela, M., Kangas, M., Korpelainen, R., & Jamsa, T. (2019). Calibration and validation of accelerometer-based activity monitors: A systematic review of machine-learning approaches. *Gait Posture*, 68, 285-299. doi:10.1016/j.gaitpost.2018.12.003
- Feng, Y., Wong, C. K., Janeja, V., Kuber, R., & Mentis, H. M. (2017). Comparison of tri-axial accelerometers step-count accuracy in slow walking conditions. *Gait & Posture*, 53, 11-16. doi:<https://doi.org/10.1016/j.gaitpost.2016.12.014>
- Fillingim, R. B., King, C. D., Ribeiro-Dasilva, M. C., Rahim-Williams, B., & Riley, J. L., 3rd. (2009). Sex, gender, and pain: a review of recent clinical and experimental findings. *J Pain*, 10(5), 447-485. doi:10.1016/j.jpain.2008.12.001
- Fitzpatrick, R., Davey, C., Buxton, M. J., & Jones, D. R. (1998). Evaluating patient-based outcome measures for use in clinical trials. *Health Technol Assess*, 2(14), i-iv, 1-74.
- Foster, N. E., Anema, J. R., Cherkin, D., Chou, R., Cohen, S. P., Gross, D. P., . . . Maher, C. G. (2018). Prevention and treatment of low back pain: evidence, challenges, and

- promising directions. *Lancet*, 391(10137), 2368-2383. doi:10.1016/s0140-6736(18)30489-6
- Garriguet, D., Tremblay, S., & Colley, R. C. (2015). Comparison of Physical Activity Adult Questionnaire results with accelerometer data. *Health Rep*, 26(7), 11-17.
- Geneen, L. J., Moore, R. A., Clarke, C., Martin, D., Colvin, L. A., & Smith, B. H. (2017). Physical activity and exercise for chronic pain in adults: an overview of Cochrane Reviews. *Cochrane Database Syst Rev*, 4, CD011279. doi:10.1002/14651858.CD011279.pub3
- Gerovasili, V., Agaku, I. T., Vardavas, C. I., & Filippidis, F. T. (2015). Levels of physical activity among adults 18-64 years old in 28 European countries. *Prev Med*, 81, 87-91. doi:10.1016/j.ypmed.2015.08.005
- Ghanei, I., Rosengren, B. E., Hasserijs, R., Nilsson, J. A., Mellstrom, D., Ohlsson, C., . . . Karlsson, M. K. (2014). The prevalence and severity of low back pain and associated symptoms in 3,009 old men. *Eur Spine J*, 23(4), 814-820. doi:10.1007/s00586-013-3139-0
- Giavarina, D. (2015). Understanding Bland Altman analysis. *Biochem Med (Zagreb)*, 25(2), 141-151. doi:10.11613/bm.2015.015
- Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. (2016). *Lancet*, 388(10053), 1545-1602. doi:10.1016/s0140-6736(16)31678-6
- Goode, A. P., Hall, K. S., Batch, B. C., Huffman, K. M., Hastings, S. N., Allen, K. D., . . . Gierisch, J. M. (2017). The Impact of Interventions that Integrate Accelerometers on Physical Activity and Weight Loss: A Systematic Review. *Ann Behav Med*, 51(1), 79-93. doi:10.1007/s12160-016-9829-1
- Grimm, E. K., Swartz, A. M., Hart, T., Miller, N. E., & Strath, S. J. (2012). Comparison of the IPAQ-Short Form and accelerometry predictions of physical activity in older adults. *J Aging Phys Act*, 20(1), 64-79.
- Grotle, M., Brox, J. I., & Vollestad, N. K. (2003). Cross-cultural adaptation of the Norwegian versions of the Roland-Morris Disability Questionnaire and the Oswestry Disability Index. *J Rehabil Med*, 35(5), 241-247.
- Grotle, M., Brox, J. I., & Vollestad, N. K. (2006). Reliability, validity and responsiveness of the fear-avoidance beliefs questionnaire: methodological aspects of the Norwegian version. *J Rehabil Med*, 38(6), 346-353. doi:10.1080/16501970600722403

- Gupta, N., Heiden, M., Mathiassen, S. E., & Holtermann, A. (2018). Is self-reported time spent sedentary and in physical activity differentially biased by age, gender, body mass index, and low-back pain? *Scand J Work Environ Health*, *44*(2), 163-170. doi:10.5271/sjweh.3693
- Gyllensten, I. C., & Bonomi, A. G. (2011). Identifying types of physical activity with a single accelerometer: evaluating laboratory-trained algorithms in daily life. *IEEE Trans Biomed Eng*, *58*(9), 2656-2663. doi:10.1109/tbme.2011.2160723
- Hale, L. A., Pal, J., & Becker, I. (2008). Measuring Free-Living Physical Activity in Adults With and Without Neurologic Dysfunction With a Triaxial Accelerometer. *Archives of Physical Medicine and Rehabilitation*, *89*(9), 1765-1771. doi:10.1016/j.apmr.2008.02.027
- Hartvigsen, J., Hancock, M. J., Kongsted, A., Louw, Q., Ferreira, M. L., Genevay, S., . . . Woolf, A. (2018). What low back pain is and why we need to pay attention. *The Lancet*, *391*(10137), 2356-2367. doi:[https://doi.org/10.1016/S0140-6736\(18\)30480-X](https://doi.org/10.1016/S0140-6736(18)30480-X)
- Helsedirektoratet. (2015). *Fysisk aktivitet og sedat tid blant voksne og eldre i Norge 2014-15 - Nasjonal kartlegging 2014-2015*. Retrieved from <https://helsedirektoratet.no/Lists/Publikasjoner/Attachments/991/Fysisk%20aktivitet%20og%20sedat%20tid%20blant%20voksne%20og%20eldre%20i%20Norge%202014-15.pdf>:
- [The Helsinki Declaration of the World Medical Association (WMA). Ethical principles of medical research involving human subjects]. (2014). *Pol Merkur Lekarski*, *36*(215), 298-301.
- Heneweer, H., Vanhees, L., & Picavet, H. S. J. (2009). Physical activity and low back pain: a U-shaped relation? *Pain*, *143*. doi:10.1016/j.pain.2008.12.033
- Heuch, I., Hagen, K., Heuch, I., Nygaard, O., & Zwart, J. A. (2010). The impact of body mass index on the prevalence of low back pain: the HUNT study. *Spine (Phila Pa 1976)*, *35*(7), 764-768. doi:10.1097/BRS.0b013e3181ba1531
- Heuch, I., Heuch, I., Hagen, K., & Zwart, J.-A. (2016). Is there a U-shaped relationship between physical activity in leisure time and risk of chronic low back pain? A follow-up in the HUNT Study. *BMC Public Health*, *16*(1), 306. doi:10.1186/s12889-016-2970-8
- Heuch, I., Heuch, I., Hagen, K., & Zwart, J. A. (2017). Physical activity level at work and risk of chronic low back pain: A follow-up in the Nord-Trøndelag Health Study. *PLoS One*, *12*(4), e0175086. doi:10.1371/journal.pone.0175086

- Hoy, D., Bain, C., Williams, G., March, L., Brooks, P., Blyth, F., . . . Buchbinder, R. (2012). A systematic review of the global prevalence of low back pain. *Arthritis Rheum*, 64(6), 2028-2037. doi:10.1002/art.34347
- Hoy, D., March, L., Brooks, P., Woolf, A., Blyth, F., Vos, T., & Buchbinder, R. (2010). Measuring the global burden of low back pain. *Best Pract Res Clin Rheumatol*, 24(2), 155-165. doi:10.1016/j.berh.2009.11.002
- Hurtig-Wennlof, A., Hagstromer, M., & Olsson, L. A. (2010). The International Physical Activity Questionnaire modified for the elderly: aspects of validity and feasibility. *Public Health Nutr*, 13(11), 1847-1854. doi:10.1017/s1368980010000157
- Ihlebaek, C., Brage, S., Natvig, B., & Bruusgaard, D. (2010). [Occurrence of musculoskeletal disorders in Norway]. *Tidsskr Nor Laegeforen*, 130(23), 2365-2368. doi:10.4045/tidsskr.09.0802
- Jago, R., Thompson, D. I., Baranowski, J. C., Cullen, K. W., & Baranowski, T. (2006). Social desirability is associated with some physical activity, psychosocial variables and sedentary behavior but not self-reported physical activity among adolescent males. *Health Education Research*, 22(3), 438-449. doi:10.1093/her/cyl107
- Jesus-Moraleida, F. R., Ferreira, P. H., Ferreira, M. L., Silva, J. P., Maher, C. G., Enthoven, W. T. M., . . . Pereira, L. S. M. (2017). Back Complaints in the Elders in Brazil and the Netherlands: a cross-sectional comparison. *Age Ageing*, 46(3), 476-481. doi:10.1093/ageing/afw230
- Jesus-Moraleida, F. R. D., Ferreira, P. H., Ferreira, M. L., Silva, J. P. D., Assis, M. G., & Pereira, L. S. M. (2018). The Brazilian Back Complaints in the Elders (Brazilian BACE) study: characteristics of Brazilian older adults with a new episode of low back pain. *Brazilian Journal of Physical Therapy*, 22(1), 55-63. doi:<https://doi.org/10.1016/j.bjpt.2017.06.018>
- Keadle, S. K., Shiroma, E. J., Kamada, M., Matthews, C. E., Harris, T. B., & Lee, I. M. (2017). Reproducibility of Accelerometer-Assessed Physical Activity and Sedentary Time. *American Journal of Preventive Medicine*, 52(4), 541-548. doi:<https://doi.org/10.1016/j.amepre.2016.11.010>
- Kerr, J., Marinac, C. R., Ellis, K., Godbole, S., Hipp, A., Glanz, K., . . . Berrigan, D. (2017). Comparison of Accelerometry Methods for Estimating Physical Activity. *Med Sci Sports Exerc*, 49(3), 617-624. doi:10.1249/mss.0000000000001124
- Kesmodel, U. S. (2018). Cross-sectional studies - what are they good for? *Acta Obstet Gynecol Scand*, 97(4), 388-393. doi:10.1111/aogs.13331

- Kinge, J. M., Knudsen, A. K., Skirbekk, V., & Vollset, S. E. (2015). Musculoskeletal disorders in Norway: prevalence of chronicity and use of primary and specialist health care services. *BMC Musculoskelet Disord*, *16*, 75. doi:10.1186/s12891-015-0536-z
- Knapik, A., Saulicz, E., & Gnat, R. (2011). Kinesiophobia - introducing a new diagnostic tool. *Journal of human kinetics*, *28*, 25-31. doi:10.2478/v10078-011-0019-8
- Kowalski, K., Rhodes, R., Naylor, P. J., Tuokko, H., & MacDonald, S. (2012). Direct and indirect measurement of physical activity in older adults: a systematic review of the literature. *Int J Behav Nutr Phys Act*, *9*, 148. doi:10.1186/1479-5868-9-148
- Laerum, E., Storheim, K., & Brox, J. I. (2007). [New clinical guidelines for low back pain]. *Tidsskr Nor Laegeforen*, *127*(20), 2706.
- Lee, I. M., Shiroma, E. J., Lobelo, F., Puska, P., Blair, S. N., & Katzmarzyk, P. T. (2012). Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. *Lancet*, *380*(9838), 219-229. doi:10.1016/s0140-6736(12)61031-9
- Lee, P. H., Macfarlane, D. J., Lam, T. H., & Stewart, S. M. (2011). Validity of the International Physical Activity Questionnaire Short Form (IPAQ-SF): a systematic review. *Int J Behav Nutr Phys Act*, *8*, 115. doi:10.1186/1479-5868-8-115
- Leopoldino, A. A., Diz, J. B., Martins, V. T., Henschke, N., Pereira, L. S., Dias, R. C., & Oliveira, V. C. (2016). Prevalence of low back pain in older Brazilians: a systematic review with meta-analysis. *Rev Bras Reumatol Engl Ed*, *56*(3), 258-269. doi:10.1016/j.rbre.2016.03.011
- Lethem, J., Slade, P. D., Troup, J. D. G., & Bentley, G. (1983). Outline of a fear-avoidance model of exaggerated pain perception—I. *Behaviour Research and Therapy*, *21*(4), 401-408. doi:[https://doi.org/10.1016/0005-7967\(83\)90009-8](https://doi.org/10.1016/0005-7967(83)90009-8)
- Levy, S. S., & Readdy, R. T. (2009). Reliability of the International Physical Activity Questionnaire in Research Settings: Last 7-Day Self-Administered Long Form. *Measurement in Physical Education and Exercise Science*, *13*(4), 191-205. doi:10.1080/10913670903260060
- Lotzke, H., Jakobsson, M., Gutke, A., Hagstromer, M., Brisby, H., Hagg, O., . . . Lundberg, M. (2018). Patients with severe low back pain exhibit a low level of physical activity before lumbar fusion surgery: a cross-sectional study. *BMC Musculoskelet Disord*, *19*(1), 365. doi:10.1186/s12891-018-2274-5

- Luque-Suarez, A., Martinez-Calderon, J., & Falla, D. (2018). Role of kinesiophobia on pain, disability and quality of life in people suffering from chronic musculoskeletal pain: a systematic review. *Br J Sports Med*. doi:10.1136/bjsports-2017-098673
- Lutz, G. K., Butzlaff, M., & Schultz-Venrath, U. (2003). Looking back on back pain: trial and error of diagnoses in the 20th century. *Spine (Phila Pa 1976)*, 28(16), 1899-1905. doi:10.1097/01.Brs.0000083365.41261.Cf
- Lyden, K., Kozey, S. L., Staudenmeyer, J. W., & Freedson, P. S. (2011). A comprehensive evaluation of commonly used accelerometer energy expenditure and MET prediction equations. *Eur J Appl Physiol*, 111(2), 187-201. doi:10.1007/s00421-010-1639-8
- Maddison, R., Jiang, Y., Hoorn, S. V., Mhurchu, C. N., Lawes, C. M., Rodgers, A., & Rush, E. (2009). Estimating energy expenditure with the RT3 triaxial accelerometer. *Res Q Exerc Sport*, 80(2), 249-256. doi:10.1080/02701367.2009.10599559
- Maher, C., Underwood, M., & Buchbinder, R. (2017). Non-specific low back pain. *The Lancet*, 389(10070), 736-747. doi:10.1016/s0140-6736(16)30970-9
- Makris, U. E., Fraenkel, L., Han, L., Leo-Summers, L., & Gill, T. M. (2014). Risk factors for restricting back pain in older persons. *J Am Med Dir Assoc*, 15(1), 62-67. doi:10.1016/j.jamda.2013.09.013
- Miguelés, J. H., Cadenas-Sanchez, C., Ekelund, U., Delisle Nystrom, C., Mora-Gonzalez, J., Lof, M., . . . Ortega, F. B. (2017). Accelerometer Data Collection and Processing Criteria to Assess Physical Activity and Other Outcomes: A Systematic Review and Practical Considerations. *Sports Med*, 47(9), 1821-1845. doi:10.1007/s40279-017-0716-0
- Mindell, J. S., Giampaoli, S., Goesswald, A., Kamtsiuris, P., Mann, C., Männistö, S., . . . Tolonen, H. (2015). Sample selection, recruitment and participation rates in health examination surveys in Europe--experience from seven national surveys. *BMC Med Res Methodol*, 15, 78. doi:10.1186/s12874-015-0072-4
- Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., . . . de Vet, H. C. W. (2010). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of Clinical Epidemiology*, 63(7), 737-745. doi:<https://doi.org/10.1016/j.jclinepi.2010.02.006>
- Montoye, A. H. K., Moore, R. W., Bowles, H. R., Korycinski, R., & Pfeiffer, K. A. (2018). Reporting accelerometer methods in physical activity intervention studies: a

- systematic review and recommendations for authors. *Br J Sports Med*, 52(23), 1507-1516. doi:10.1136/bjsports-2015-095947
- Motl, R. W., McAuley, E., & DiStefano, C. (2005). Is social desirability associated with self-reported physical activity? *Prev Med*, 40(6), 735-739. doi:<https://doi.org/10.1016/j.ypmed.2004.09.016>
- Naugle, K. M., Ohlman, T., Naugle, K. E., Riley, Z. A., & Keith, N. R. (2017). Physical activity behavior predicts endogenous pain modulation in older adults. *Pain*, 158(3), 383-390. doi:10.1097/j.pain.0000000000000769
- Ng, S. W., Howard, A.-G., Wang, H. J., Su, C., & Zhang, B. (2014). The physical activity transition among adults in China: 1991–2011. *15*(S1), 27-36. doi:doi:10.1111/obr.12127
- Ng, S. W., & Popkin, B. (2012). Time Use and Physical Activity: A Shift Away from Movement across the Globe. *Obesity Reviews*, 13(8), 659-680. doi:10.1111/j.1467-789X.2011.00982.x
- O'Neill, B., McDonough, S. M., Wilson, J. J., Bradbury, I., Hayes, K., Kirk, A., . . . Tully, M. A. (2017). Comparing accelerometer, pedometer and a questionnaire for measuring physical activity in bronchiectasis: a validity and feasibility study? *Respir Res*, 18(1), 16. doi:10.1186/s12931-016-0497-2
- O'Sullivan, P. B., Caneiro, J. P., O'Keeffe, M., Smith, A., Dankaerts, W., Fersum, K., & O'Sullivan, K. (2018). Cognitive Functional Therapy: An Integrated Behavioral Approach for the Targeted Management of Disabling Low Back Pain. *Phys Ther*, 98(5), 408-423. doi:10.1093/ptj/pzy022
- Palacios-Cena, D., Alonso-Blanco, C., Hernandez-Barrera, V., Carrasco-Garrido, P., Jimenez-Garcia, R., & Fernandez-de-las-Penas, C. (2015). Prevalence of neck and low back pain in community-dwelling adults in Spain: an updated population-based national study (2009/10-2011/12). *Eur Spine J*, 24(3), 482-492. doi:10.1007/s00586-014-3567-5
- Paulhus, D. L. (1991). CHAPTER 2 - Measurement and Control of Response Bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of Personality and Social Psychological Attitudes* (pp. 17-59): Academic Press.
- Phan, D. H., Bonnet, S., Guillemaud, R., Castelli, E., & Pham Thi, N. Y. (2008). Estimation of respiratory waveform and heart rate using an accelerometer. *Conf Proc IEEE Eng Med Biol Soc*, 2008, 4916-4919. doi:10.1109/iembs.2008.4650316

- Qaseem, A., Wilt, T. J., McLean, R. M., Forcica, M. A., & Clinical Guidelines Committee of the American College of P. (2017). Noninvasive Treatments for Acute, Subacute, and Chronic Low Back Pain: A Clinical Practice Guideline From the American College of Physicians. *Ann Intern Med*, *166*(7), 514-530. doi:10.7326/M16-2367
- Roland, M., & Morris, R. (1983). A study of the natural history of back pain. Part I: development of a reliable and sensitive measure of disability in low-back pain. *Spine (Phila Pa 1976)*, *8*(2), 141-144.
- Rowlands, A. V., Mirkes, E. M., Yates, T., Clemes, S., Davies, M., Khunti, K., & Edwardson, C. L. (2018). Accelerometer-assessed Physical Activity in Epidemiology: Are Monitors Equivalent? *Med Sci Sports Exerc*, *50*(2), 257-265. doi:10.1249/mss.0000000000001435
- Sanda, B., Vistad, I., Haakstad, L. A. H., Berntsen, S., Sagedal, L. R., Lohne-Seiler, H., & Torstveit, M. K. (2017). Reliability and concurrent validity of the International Physical Activity Questionnaire short form among pregnant women. *BMC Sports Science, Medicine and Rehabilitation*, *9*(1). doi:10.1186/s13102-017-0070-4
- Scheele, J., Luijsterburg, P. A. J., Ferreira, M. L., Maher, C. G., Pereira, L., Peul, W. C., . . . Koes, B. W. (2011). Back Complaints in the Elders (BACE); design of cohort studies in primary care: an international consortium. *BMC Musculoskeletal Disorders*, *12*(1), 193. doi:10.1186/1471-2474-12-193
- Schmal, H., Holsgaard-Larsen, A., Izadpanah, K., Brond, J. C., Madsen, C. F., & Lauritsen, J. (2018). Validation of Activity Tracking Procedures in Elderly Patients after Operative Treatment of Proximal Femur Fractures. *Rehabil Res Pract*, *2018*, 3521271. doi:10.1155/2018/3521271
- Schneller, M. B., Bentsen, P., Nielsen, G., Brond, J. C., Ried-Larsen, M., Mygind, E., & Schipperijn, J. (2017). Measuring Children's Physical Activity: Compliance Using Skin-Taped Accelerometers. *Med Sci Sports Exerc*, *49*(6), 1261-1269. doi:10.1249/mss.0000000000001222
- Scott, J. J., Rowlands, A. V., Cliff, D. P., Morgan, P. J., Plotnikoff, R. C., & Lubans, D. R. (2017). Comparability and feasibility of wrist- and hip-worn accelerometers in free-living adolescents. *J Sci Med Sport*, *20*(12), 1101-1106. doi:10.1016/j.jsams.2017.04.017
- Shipton, E. A. (2018). Physical Therapy Approaches in the Treatment of Low Back Pain. *Pain and Therapy*. doi:10.1007/s40122-018-0105-x

- Shiri, R., & Falah-Hassani, K. (2017). Does leisure time physical activity protect against low back pain? Systematic review and meta-analysis of 36 prospective cohort studies. *Br J Sports Med*, *51*(19), 1410-1418. doi:10.1136/bjsports-2016-097352
- Silsbury, Z., Goldsmith, R., & Rushton, A. (2015). Systematic review of the measurement properties of self-report physical activity questionnaires in healthy adult populations. *BMJ Open*, *5*(9), e008430. doi:10.1136/bmjopen-2015-008430
- Skender, S., Ose, J., Chang-Claude, J., Paskow, M., Bruhmann, B., Siegel, E. M., . . . Ulrich, C. M. (2016). Accelerometry and physical activity questionnaires - a systematic review. *BMC Public Health*, *16*, 515. doi:10.1186/s12889-016-3172-0
- Steene-Johannessen, J., Anderssen, S. A., van der Ploeg, H. P., Hendriksen, I. J., Donnelly, A. E., Brage, S., & Ekelund, U. (2016). Are Self-report Measures Able to Define Individuals as Physically Active or Inactive? *Med Sci Sports Exerc*, *48*(2), 235-244. doi:10.1249/mss.0000000000000760
- Steffens, D., Maher, C. G., Pereira, L. S., Stevens, M. L., Oliveira, V. C., Chapple, M., . . . Hancock, M. J. (2016). Prevention of Low Back Pain: A Systematic Review and Meta-analysis. *JAMA Intern Med*, *176*(2), 199-208. doi:10.1001/jamainternmed.2015.7431
- Stenberg, G., Fjellman-Wiklund, A., & Ahlgren, C. (2014). 'I am afraid to make the damage worse'--fear of engaging in physical activity among patients with neck or back pain--a gender perspective. *Scand J Caring Sci*, *28*(1), 146-154. doi:10.1111/scs.12043
- Stewart, T., Narayanan, A., Hedayatrad, L., Neville, J., Mackay, L., & Duncan, S. (2018). A Dual-Accelerometer System for Classifying Physical Activity in Children and Adults. *Med Sci Sports Exerc*. doi:10.1249/mss.0000000000001717
- Straczkiwicz, M., Urbanek, J. K., Fadel, W. F., Crainiceanu, C. M., & Harezlak, J. (2016). Automatic car driving detection using raw accelerometry data. *Physiol Meas*, *37*(10), 1757-1769. doi:10.1088/0967-3334/37/10/1757
- Sylvia, L. G., Bernstein, E. E., Hubbard, J. L., Keating, L., & Anderson, E. J. (2014). A Practical Guide to Measuring Physical Activity. *Journal of the Academy of Nutrition and Dietetics*, *114*(2), 199-208. doi:10.1016/j.jand.2013.09.018
- Søgaard, K., & Sjøgaard, G. (2017). Physical Activity as Cause and Cure of Muscular Pain: Evidence of Underlying Mechanisms. *Exercise and sport sciences reviews*, *45*(3), 136-145. doi:10.1249/JES.0000000000000112
- Terwee, C. B., Bot, S. D. M., de Boer, M. R., van der Windt, D. A. W. M., Knol, D. L., Dekker, J., . . . de Vet, H. C. W. (2007). Quality criteria were proposed for

- measurement properties of health status questionnaires. *Journal of Clinical Epidemiology*, 60(1), 34-42. doi:<https://doi.org/10.1016/j.jclinepi.2006.03.012>
- Tomioka, K., Iwamoto, J., Saeki, K., & Okamoto, N. (2011). Reliability and validity of the International Physical Activity Questionnaire (IPAQ) in elderly adults: the Fujiwarakyo Study. *J Epidemiol*, 21(6), 459-465.
- Troiano, R. P. (2009). Can there be a single best measure of reported physical activity? *The American journal of clinical nutrition*, 89(3), 736-737. doi:10.3945/ajcn.2008.27461
- Troiano, R. P., McClain, J. J., Brychta, R. J., & Chen, K. Y. (2014). Evolution of accelerometer methods for physical activity research. *Br J Sports Med*, 48(13), 1019-1023. doi:10.1136/bjsports-2014-093546
- Trost, S. G., & O'Neil, M. (2014). Clinical use of objective measures of physical activity. *Br J Sports Med*, 48(3), 178-181. doi:10.1136/bjsports-2013-093173
- Tulchinsky, T. H., Varavikova, E. A., & Varavikova, E. A. (2000). *The New Public Health : An Introduction for the 21st Century*. San Diego, UNITED STATES: Elsevier Science & Technology.
- van Hees, V. T., Renström, F., Wright, A., Gradmark, A., Catt, M., Chen, K. Y., . . . Franks, P. W. (2011). Estimation of Daily Energy Expenditure in Pregnant and Non-Pregnant Women Using a Wrist-Worn Tri-Axial Accelerometer. *PLoS One*, 6(7), e22922. doi:10.1371/journal.pone.0022922
- Van Holle, V., De Bourdeaudhuij, I., Deforche, B., Van Cauwenberg, J., & Van Dyck, D. (2015). Assessment of physical activity in older Belgian adults: validity and reliability of an adapted interview version of the long International Physical Activity Questionnaire (IPAQ-L). *BMC Public Health*, 15, 433. doi:10.1186/s12889-015-1785-3
- van Poppel, M. N., Chinapaw, M. J., Mokkink, L. B., van Mechelen, W., & Terwee, C. B. (2010). Physical activity questionnaires for adults: a systematic review of measurement properties. *Sports Med*, 40(7), 565-600. doi:10.2165/11531930-000000000-00000
- van Weering, M. G., Vollenbroek-Hutten, M. M., & Hermens, H. J. (2011). The relationship between objectively and subjectively measured activity levels in people with chronic low back pain. *Clin Rehabil*, 25(3), 256-263. doi:10.1177/0269215510380828
- van Weering, M. G., Vollenbroek-Hutten, M. M., Tonis, T. M., & Hermens, H. J. (2009). Daily physical activities in chronic lower back pain patients assessed with accelerometry. *Eur J Pain*, 13(6), 649-654. doi:10.1016/j.ejpain.2008.07.005

- Varela Mato, V., Yates, T., Stensel, D., Biddle, S., & Clemes, S. A. (2017). Concurrent Validity of Actigraph-Determined Sedentary Time Against the ActiVpal Under Free-Living Conditions in a Sample of Bus Drivers. *Measurement in Physical Education and Exercise Science*, 21(4), 212-222. doi:10.1080/1091367X.2017.1335204
- Vibe Fersum, K., O'Sullivan, P., Skouen, J. S., Smith, A., & Kvale, A. (2013). Efficacy of classification-based cognitive functional therapy in patients with non-specific chronic low back pain: a randomized controlled trial. *Eur J Pain*, 17(6), 916-928. doi:10.1002/j.1532-2149.2012.00252.x
- Vos, T., Abajobir, A. A., Abate, K. H., Abbafati, C., Abbas, K. M., Abd-Allah, F., . . . Murray, C. J. L. (2017). Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet*, 390(10100), 1211-1259. doi:10.1016/s0140-6736(17)32154-2
- Waddell, G., Newton, M., Henderson, I., Somerville, D., & Main, C. J. (1993). A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. *Pain*, 52(2), 157-168.
- Weiner, D. K., Haggerty, C. L., Kritchevsky, S. B., Harris, T., Simonsick, E. M., Nevitt, M., & Newman, A. (2003). How does low back pain impact physical function in independent, well-functioning older adults? Evidence from the Health ABC Cohort and implications for the future. *Pain Med*, 4(4), 311-320.
- Wertli, M. M., Rasmussen-Barr, E., Held, U., Weiser, S., Bachmann, L. M., & Brunner, F. (2014). Fear-avoidance beliefs-a moderator of treatment efficacy in patients with low back pain: a systematic review. *Spine J*, 14(11), 2658-2678. doi:10.1016/j.spinee.2014.02.033
- Westerterp, K. R. (2009). Assessment of physical activity: a critical appraisal. *Eur J Appl Physiol*, 105(6), 823-828. doi:10.1007/s00421-009-1000-2
- Westerterp, K. R. (2017). Doubly labelled water assessment of energy expenditure: principle, practice, and promise. *Eur J Appl Physiol*, 117(7), 1277-1285. doi:10.1007/s00421-017-3641-x
- White, T., Westgate, K., Hollidge, S., Venables, M., Olivier, P., Wareham, N., & Brage, S. (2019). Estimating energy expenditure from wrist and thigh accelerometry in free-living adults: a doubly labelled water study. *Int J Obes (Lond)*. doi:10.1038/s41366-019-0352-x

WHO. Physical activity and Older Adults Retrieved from

https://www.who.int/dietphysicalactivity/factsheet_olderadults/en/

- Wijndaele, K., Westgate, K., Stephens, S. K., Blair, S. N., Bull, F. C., Chastin, S. F., . . . Healy, G. N. (2015). Utilization and Harmonization of Adult Accelerometry Data: Review and Expert Consensus. *Med Sci Sports Exerc*, *47*(10), 2129-2139. doi:10.1249/mss.0000000000000661
- Willettts, M., Hollowell, S., Aslett, L., Holmes, C., & Doherty, A. (2018). Statistical machine learning of sleep and physical activity phenotypes from sensor data in 96,220 UK Biobank participants. *Sci Rep*, *8*(1), 7961. doi:10.1038/s41598-018-26174-1
- Williams, B., Entwistle, V., Haddow, G., & Wells, M. (2008). Promoting research participation: why not advertise altruism? *Soc Sci Med*, *66*(7), 1451-1456. doi:10.1016/j.socscimed.2007.12.013
- Winckers, A. N., Mackenbach, J. D., Compernelle, S., Nicolaou, M., van der Ploeg, H. P., De Bourdeaudhuij, I., . . . Lakerveld, J. (2015). Educational differences in the validity of self-reported physical activity. *BMC Public Health*, *15*, 1299. doi:10.1186/s12889-015-2656-7
- Wong, A. Y. L., Karppinen, J., & Samartzis, D. (2017). Low back pain in older adults: risk factors, management options and future directions. *Scoliosis and Spinal Disorders*, *12*, 14. doi:10.1186/s13013-017-0121-3