

Evaluering av e-læringsressursen "Klinisk undersøkelse av bevegelsesapparatet"

<http://meddev.uio.no/elaring/fag/ortopedi/index.shtml>

Alexander Øvestad og Caroline Kooy Tveiten



Prosjektoppgave ved Det medisinske fakultet

UNIVERSITETET I OSLO

22.01.2018

I Abstract

Purpose: to evaluate the improvement of the new internet resource "The Musculoskeletal System" compared to the 10 year old system, and to review the effectiveness and usage of e-learning resources for the purpose to assess future requirements.

Material and Method: A crossover design was used. Our participants were 8th and 9th semester medical students. They were recruited by a common announcement sent to their respective class Facebook page, and by direct message on Facebook in alphabetical order from the members list on said page. Twenty students participated and were randomly allocated to two groups, containing ten students each. One group studied the old system, the other studied the new system. Both groups then tested their performance in our quiz, and filled in a questionnaire (about background, comparison, System Usability Scale and qualitative questions including time-use). A wash out period consisting of minimum two weeks followed, and the interventions were switched for the two groups, who then repeated the quiz and filled in the remainder of the questionnaire. The study lasted from June 2017 to January 2018. The performance of each participant was rated 0-100 for each category, and 0-500 in total for all five categories tested. The quantitative questions in the questionnaire were answered on a 0-5 scale.

Results: The quiz showed a positive but statistically insignificant difference in performance in favor of the new system compared to the old system. The questionnaire showed statistically significant differences for all questions comparing the two systems, also in favor of the new system. It also showed a statistically significant difference in time-use, where the participants spent less time on the new system.

Conclusion: Based on the questionnaire, the new system is an improvement compared to the old one. The participants are positive with regards to future use of e-learning resources.

II Innholdsfortegnelse

I	Abstract	s. 1
II	Innholdsfortegnelse	s. 2
1	Innledning	s. 3
	1.1 Bakgrunn for valg av oppgave	s. 3
	1.2 Valg av oppgave	s. 6
	1.3 Problemstilling	s. 7
	1.3.1 Definisjon av begreper i problemstillingen	s. 7
	1.3.2 Avgrensning av problemstillingen	s. 8
2	Materiale	s. 9
3	Metode	s. 10
	3.1 Crossover-studie	s. 10
	3.2 Spørreundersøkelse	s. 11
	3.3 Fremgangsmåte ved datainnsamling	s. 12
	3.3.1 Utvalg og gjennomføring	s. 12
	3.3.2 Anonymisering	s. 13
	3.3.3 Innholdet i undersøkelsene	s. 13
	3.3.4 Evaluering og metodekritikk	s. 15
4	Resultater	s. 17
	4.1 Innhenting og svarprosent	s. 17
	4.2 Quiz	s. 18
	4.2 Spørreundersøkelse	s. 24
5	Diskusjon	s. 34
	5.1 Materiale	s. 34
	5.2 Studiedesign	s. 35
	5.3 Resultater	s. 36
	5.3.1 Innhenting og Svarprosent	s. 36
	5.3.2 Quiz	s. 38
	5.3.3 Spørreundersøkelse	s. 41
6	Konklusjon	s. 45
7	Referanser	s. 46
III	Vedlegg	s. 48

1 Innledning

Formålet med denne prosjektoppgaven er å evaluere om arbeidet vårt med å lage en oppdatert versjon av e-læringsressursen "Bevegelsesapparatet" har lyktes. En pilot-brukertest ble utført for å undersøke om vi var på rett vei i revisjonsarbeidet (kne e-læringen ble brukt i testen), for å kunne identifisere brukerfeil i designet og for å imøtekomme andre ønsker og samle inn nye ideer for dermed å kunne gjøre de andre delene av e-læringsressursen enda bedre. Vi vil også belyse e-læringsprogrammets rolle i undervisningssammenheng, for å finne ut av om dette er en hensiktsmessig bruk av ressurser for det medisinske fakultet ved Universitetet i Oslo i fremtiden. Vi vil til slutt undersøke om den nye ressursen ivaretar de nye læreplanmålene i den nye reformen av profesjonsstudiet i medisin (Oslo 2014).

1.1 Bakgrunn for valg av oppgave

Profesjonsstudiet i medisin er et studium som strekker seg over seks år, med en fastlagt undervisningsplan. Tidligere har studiet vært inndelt med ulike fagemner fordelt over 12 semestre. Fagområdet "Bevegelsesapparatet", bestående av anatomi og klinisk undersøkelse av muskel- og skjelettsystemet, var lagt inn under den andre halvdel av fjerde semester (tidligere kalt 4B). Grunntanken i den pedagogiske reformen à Oslo 96 var å fjerne skillet mellom en tidligere strengt adskilt teori og praksis. Studentene ble her fordelt i grupper på ulike avdelinger på sykehusene Ahus, Rikshospitalet, Ullevål og Diakonhjemmet. På sykehusene foregikk den pasientrelaterte undervisningen i smågrupper, parallelt med at det ble undervist i "funksjonsundersøkelse" på universitetet.

Tanken med mer pasientkontakt og en mer klinisk rettet undervisning, kombinert med læringen av anatomi og fysiologi, skapte imidlertid mye frustrasjon på 4B-semesteret. Grunnen til dette var at studentene lærte forskjellige undersøkelsesmetoder avhengig av hvilken avdeling og hvilket sykehus de var på. I tillegg ble "funksjonsundersøkelse" på Universitetet i hovedsak undervist av anatomer, som var mye mindre klinisk rettet. Dette førte til en stor variasjon av mange ulike undersøkelsesmetoder og fremgangsmåter - og dermed også et stort behov for systematikk. I den forbindelse ble det fremlagt et ønske om en

felles kunnskapsplattform for både studenter og undervisere, som etterhvert tok form av et IKT-basert læringsprogram. Den utgikk fra et samarbeid mellom de ulike faggruppene ortopedi, fysikalsk medisin og anatomi. Målet var en konsensus for læreplanmålene mellom de ulike spesialitetene, formidlet ved hjelp av skriftlig materiale og videofilmer i en lett tilgjengelig kunnskapskilde.

Som en følge av dette ble internettressursen "Klinisk undersøkelse av bevegelsesapparatet" ferdigstilt i 2004 (1), laget av de fire medisinstudentene Jill Gunnel Eikrem Bruaset, Ida Kristin Haugen, Trine-Lise Kristoffersen og Trine Lise Svinningen Strand. Arbeidet ble utført under veiledning av Eric Rinvik (anatom), Niels Gunnar Juel (fysikalskmedisiner) og Lars Engebretsen (ortoped). Disse e-læringsleksjonene har vært til stor glede for mange studenter i ettertid, også for undertegnende, da vi gikk 4B semesteret i 2015. Forskning og rutiner for undersøkelsesmetoder har imidlertid endret seg over det siste tiåret, og kravene til teknologisk kvalitet har økt betraktelig. Da medisinstudiet på nytt skulle revideres i 2014, ble det avgjort at en oppdatering av alt det tidligere materialet var en nødvendighet. Tanken var at arbeidet også denne gangen skulle utføres av studenter under veiledning, noe som fanget vår interesse da ideen ble presentert under en forelesning med Lars Engebretsen. Vi hadde selv lagt merke til potensialet i den gamle ressursen, men samtidig opplevd et ønske om en mer relevant, presis og oppdatert versjon.

Den reviderte studieplanen i Oslo 2014 ble innført i profesjonsstudiet i medisin i løpet av 2017. Den innebærer blant annet en overgang fra semestre til et modulsystem bestående av 8 moduler. Kunnskap om bevegelsesapparatet skal ikke lenger være på 4B-semesteret, men er nå fordelt på modul 3 og 8. I modul 3 skal studentene tilegne seg grunnkunnskaper om bevegelsesapparatets normale struktur og funksjon, samt opplæres i relevant undersøkelsesmetodikk. Modul 8 dreier seg i hovedsak om kunnskap i patologien, å kunne bruke relevante undersøkelser videre i det diagnostiske arbeidet, samt videre behandling av pasientens problem (er). (Vedlegg 1: Læreplanmål for modul 3 og 8.)

Tanken bak en ny e-læringsressurs er å igjen skape et skarpere skille mellom anatomi og klinikk. Anatomi dreier seg om et dynamisk kunnskapsområde og er et universelt fagfelt. Å skape en god internettside for anatomien vil kreve enorme ressurser. Det vil neppe være hensiktsmessig for oss å lage en slik nettside, med tanke på at det allerede eksisterer utallige gode informasjonskilder om anatomi fra tidligere arbeid, blant annet i form av illustrerende

atlas, app'er og videoer på YouTube. Behovet for en god *klinisk* rettet kunnskapskilde er derimot svært relevant, og skal derfor være hovedfokus i den nye ressursen. Vi samlet informasjon om hva som fungerte henholdsvis bra og dårlig i den eksisterende nettressursen, basert på vårt eget og våre medstudenters inntrykk, etter å ha gjennomført vårt 4B-semester. Denne informasjonen ble så grunnlaget for et revisjonsprosjekt med det formål å ta vare på det som fungerte bra, og å forbedre det som ble oppfattet som dårlig. Dette har vært rammen for det revisjonsarbeidet vi har gjort. Her er et utkast av vår liste med revisjonsforbedringene:

Det som fungerte godt i gammel ressurs (og som vi vil forsøke å ivareta i den nye):

- Ressursen ivaretar (den gamle) studieplanens læringsmål
- Den inneholder mye bra og nyttig stoff
- Den er tydelig basert på et grundig og detaljert arbeide
- Den ivaretar både undervisning i klinikk og FU, og fungerer som en felles plattform (i den grad det er mulig)
- Ressursen har en logisk og oversiktlig inndeling

Det som fungerte mindre bra i gammel ressurs:

- Den er for detaljert: det er mye nyttig info, men vanskelig å skille det mest relevante fra "detalj-kunnskap"
- Ressursen har en blanding av anatomi og klinisk fokus, som gjør det vanskelig å vite hva som faktisk skal gjøres i klinikken. (Hensikten var jo initialt å vise hvilken del av anatomien som var viktig for klinikken- men det ble i stedet altfor mye stoff og detaljer)
- Den er ikke bygd opp slik vi lærer å ta opp en journal i dag
- Det er få illustrasjoner, kombinert med til tider veldig mye tekst
- Den er lite interaktiv, og dermed litt kjedelig og uinspirerende
- Det er altfor mange videoer (og det er vanskelig å skille viktigheten og relevansen av dem)
- Treige filmer (både i teknisk kvalitet og i innhold)
- Videoene virker utdatert i layout, slik at man stiller seg spørsmål om troverdighet i forhold til om kunnskapen er oppdatert.

- Nettsiden har for mange ”trykk”, slik at det er vanskelig å få oversikt over informasjonsmengden, og dermed vanskelig å legge opp en lesestrategi for å komme seg igjennom alt stoffet.
- Ressursen ivaretar ikke helt de nye læreplanmålene, ettersom ressursen er ment for den gamle reformen.

1.2 Valg av oppgave

Det overordnede målet med oppdateringen av e-læringsressursen "Klinisk undersøkelse av bevegelsesapparatet" er å lage en oppdatert, troverdig og lærerik ressurs som både fungerer godt i klinikken, men også rent teknisk. Den skal være så god i både sitt innhold og sin "layout" at den vil bli brukt av våre medstudenter; spesielt på medisinstudiet men også gjerne i senere turnustjeneste.

Til tross for et ønske om (og nå delvis gjennomførte) forbedringer, har det ikke blitt dokumentert noen effekt av bruken av e-læringsressursen. Vi anser det som viktig å undersøke studentenes bruk og erfaring med denne nettressursen, både med tanke på å forbedre den mest mulig her og nå, men også for å kartlegge nødvendigheten av fremtidig fokus på å prioritere og lage slike nettbaserte kunnskapskilder. Kunnskap om dette vil være til hjelp for å vurdere om det er hensiktsmessig å legge ned tid og penger i liknende ressurser på andre fagområder også. Internettbasert læring sammenliknet med *ingen* intervensjon, er assosiert med store positive effekter. Men når man sammenligner læringseffektiviteten av internettbasert læring med andre tradisjonelle læringsmetoder, er det stor heterogenitet i resultatene (2).

Det er imidlertid svært vanskelig å gjennomføre studier på effektiviteten og nytten av internettbaserte læringsressurser (IBL), fordi disse kan variere enormt i både oppbygning, innhold og kvalitet. Dette er gjennomgående når man leser om forskning gjort på dette feltet. Generelt er det vanskelig å studere slike læringsverktøy fordi de for det første har så varierende utforming, som for eksempel ”chat”, gruppearbeid, tester og brukergrensesnitt. For det andre gjelder det faktum at de brukes i veldig forskjellige settinger, som for eksempel for ulike typer studenter, som supplerende undervisningsmetoder, koordinering o.l. To studier

om samme problemstilling kan derfor komme frem til stikk motsatt resultat, og det er vanskelig å overføre resultatene fra et forsøk til et annet. Internettbasert læring kan derfor ikke ansees som én enhet, fordi det finnes så mange ulike konfigurasjoner og metoder å utforme dem på. Det viser seg å være så stor variasjon at det kan være vanskelig å sammenligne dem (3).

Vår hypotese er at jo mer oversiktlig, visuell og interaktiv en ressurs er, jo større og mer effektivt vil læringsutbyttet også være. En oversiktsartikkel fra 2010 som ser på den relative effekten av høy versus lav interaktivitet, feedback, diskusjon og spirallæring i internettbasert læring for helseutdanning, understøtter dette. Til tross for heterogeniteten mellom ulike studier som ble sammenliknet, så var det tydelig at interaktivitet, praktiske øvelser, repetisjon og feedback var assosiert med økt læringsutbytte (4). Vår nye e-læringsressurs er basert på nettopp dette; den inkluderer videofilmer, pasientkasus og quizer, der studentene mottar en umiddelbar tilbakemelding med både resultat og en forklaring på hvorfor avgitt svarrespons var rett eller gal.

For å kunne teste forbedringene vi hadde gjort på best mulig måte, og for å få besvart de spørsmålene vi hadde om læringseffektiviteten, valgte vi å gjennomføre vårt forskningsprosjekt som en crossover-studie med en etterfølgende spørreundersøkelse, der vi ville sammenlikne den nye ressursen med den fra 2004.

1.3 Problemstilling

“Har vi klart å forbedre det som trengte forbedring, og å ivareta det som var bra, i arbeidet med oppdateringen av e-læringsressursen "Klinisk undersøkelse av bevegelsesapparatet"?"

1.3.1 Definisjon av begreper i problemstillingen

"E-læringsressursen" henviser til det internettbaserte læringsprogrammet om klinisk undersøkelse av bevegelsesapparatet. Den består av åtte web-leksjoner utarbeidet i 2004 for Universitetet i Oslo. Det er et internettbasert læringsprogram bestående av anatomi og

undersøkelse av de åtte leddene skulder, albue, hånd, rygg, nakke, hofta, kne og ankel/fot. Vi har her valgt å kun bruke leksjonen som omhandler *kneet* som en prototype.

Med "oppdatering" mener vi at vårt arbeid har tatt utgangspunkt i den versjonen av "Klinisk undersøkelse av bevegelsesapparatet" som ble utført i 2004, da web-leksjonene ble laget for aller første gang.

Med ordene "forbedring" og "bra" har vi tatt utgangspunkt i en rekke punkter som vi kom frem til ved evalueringen av web-leksjonene à 2004. Disse er videre utdypet i diskusjonen nedenunder.

1.3.2 Avgrensning av problemstillingen

Vi har valgt å gjennomføre denne undersøkelsen på grunnlag av kun én av de åtte web-leksjonene, og den er begrenset til et utvalg av studenter på to kull. Årsaken til dette er i hovedsak tidsmessige begrensninger og vil utdypes videre under metodekapittelet.

I denne undersøkelsen ønsker vi også å danne oss et bilde av bruken av internett som en kilde til å få dekket læreplanmålene på medisinstudiet i Oslo. Store økonomiske ressurser og mye tidkrevende arbeidskraft ligger bak et slikt prosjekt. Dette går litt utover vår problemstilling, men vi har allikevel valgt å ta med noen spørsmål angående dette fordi det er svært relevant for å vurdere nytten av å lage slike ressurser i fremtiden.

2 Materiale

Redigeringsarbeidet av den delen av e-læringsressursen som omhandler *kneet* ble utført i et samarbeid mellom medisinstudentene Hannah Ugletveit Jahr og Caroline Kooy Tveiten. Per Grøttum (Professor ved Seksjon for medisinsk informatikk) fungerte som teknisk veileder, og Lars Engebretsen (professor og ortoped) som faglig veileder. De tilhørende videofilmene ble filmet og redigert av Morten Skoglund (Seksjon for medisinsk informatikk) med Lars Engebretsen som undersøkende lege, og Caroline Kooy Tveiten som "modell". Delemnet om *kneet* ble ferdigstilt i 2016, og web-leksjonen skulle videre fungere som prototype for de andre leddene. Av flere grunner (som nevnt i innledningen), valgte vi å bruke kun ett ledd når vi skulle utføre dette prosjektet, og vi valgte da kneleddet som læringsobjekt for deltakerne i denne undersøkelsen. Vi rekrutterte 20 medisinstudenter fra kullene H13 og V14, som frivillig valgte å delta i prosjektet vårt.

Link til e-læringsressurs (2004): <http://meddev.uio.no/elaring/fag/anatomi/index.shtml>

Link til ny e-læringsressurs (2016): <http://meddev.uio.no/elaring/fag/ortopedi/index.shtml>

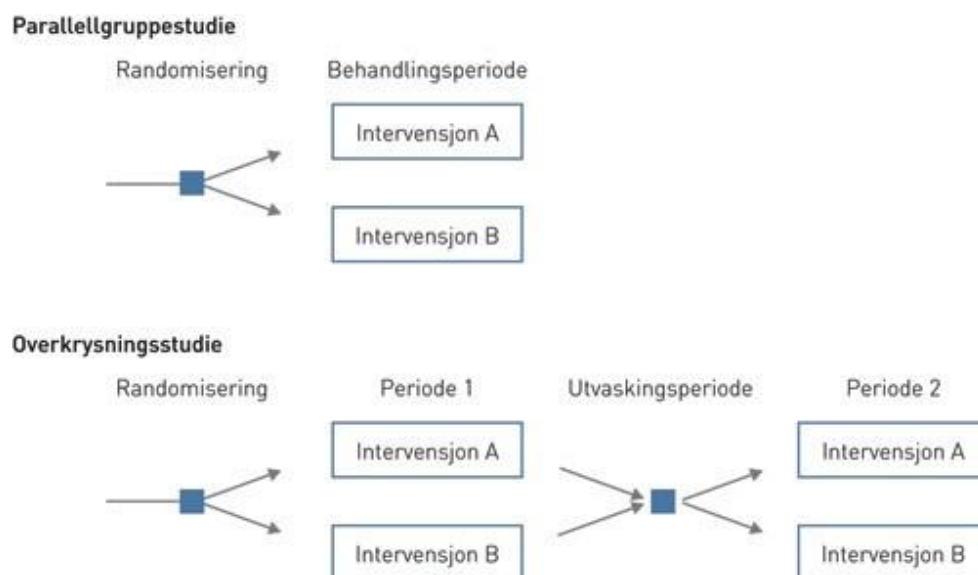
En rekke punkter om ønsker om forbedringer av den gamle ressursen ble utarbeidet av oss, og fungerte som substans for spørsmålene i spørreundersøkelsen, altså grunnlaget vårt for prosjektoppgaven (disse er videre utdypet i diskusjonen).

3 Metode

Vi har valgt å sammenligne to systemer ved hjelp av en randomisert strukturert crossover- og spørreskjema studie. Her presenteres de aktuelle studiedesign, samt fremgangsmåten for gjennomføringen av informasjonsinnhenting, og analysen av den.

3.1 Crossover-studie

En crossover-studie, også kalt overkrysningsstudie (5), skiller seg fra randomiserte forsøk, som er gullstandarden for å avgjøre om en intervensjon er bedre enn en annen. Ved et randomisert forsøk sammenlikner man to eller flere grupper som får hver sin intervensjon ved at disse gruppene følges parallelt, og det kan derfor også kalles en parallellestudie. Dersom en studie handler om å observere noe som er relativt stabilt og varer over tid, kan det imidlertid være hensiktsmessig med et annet studiedesign, der deltakerne først kan få en type intervensjon, og deretter en annen. De to intervensjonsperiodene må skilles av en tidsperiode, en såkalt "utvaskingsperiode", som er lang nok til å sikre at man ikke får noen overføringseffekt mellom de to intervensjonene. Hver deltaker fungerer her som sin egen kontroll. En crossover-studie krever færre pasienter enn et randomisert forsøk, men det er mer tidkrevende fordi man gir intervensjonene etter hverandre i serie.



Figur 1

3.2 Spørreundersøkelse

En spørreundersøkelse er definert som betegnelsen på en gruppe metoder for relativt strukturert datainnsamling ved hjelp av spørsmål og svar. Det er en indirekte metode som er lett å administrere, har lite kostnader og som sikrer anonymitet (6). En spørreundersøkelse kan klassifiseres som både en kvalitativ og en kvantitativ metode, avhengig av spørsmålenes utforming. Typisk vil "lukkede spørsmål" med multiple svaralternativer egne seg godt til kvantitativ analyse. Spørsmålene her vil gjøre det lett å sammenlikne og å kode svarene, men en utfordring er at de kan være vanskelige å formulere fordi de må være svært presise for å unngå misforståelser. En spørreundersøkelse med mer "åpne spørsmål" vil være lettere å formulere og eventuelle misforståelser vil kunne oppdages, men svarene her er derimot vanskeligere å sammenlikne, og man kan feiltolke kodingen. Uavhengig om man bruker den til kvantitative eller kvalitative formål, så er spørreundersøkelse en metode som er billig, lett gjennomførbar og som raskt kan innhente store mengder informasjon med en ofte høy grad av objektivitet. Elementer som man må ta i betraktning er at det kan være fare for "tilfeldige svar" dersom deltakeren ikke leser spørsmålene nøye nok, eller at han/hun ikke får uttrykt sine meninger fullt ut som en følge av irrelevante eller dårlig formulerte spørsmål. Det er viktig at spørreundersøkelsen også har et språk og en "layout" som motiverer deltakerne til å svare. Spørreskjemaet bør innledes med faktaorienterte oppvarmingsspørsmål, etterfulgt av mer kontroversielle holdningsspørsmål og deretter avrundes med ukontroversielle spørsmål (7).

3.3 Fremgangsmåte ved datainnsamling

3.3.1 Utvalg og gjennomføring

Før pilotprosjektet vårt kunne starte, måtte studentene introduseres for både den gamle og den nye ressursen. De ble forklart hvor de kunne finne fram og gitt tilgang til den nye ressursen (ettersom den ikke offisielt ligger ute) og de ble videre gitt en kort innføring i hvordan ressursene skulle brukes. De ble også gjort oppmerksomme på viktigheten av deltagelsen deres. Videre ble de informert om rettigheten til å trekke seg når som helst i løpet av studieperioden, uten å måtte forklare hvorfor, om at full anonymitet ville bli beholdt, og om hvordan resten av prosjektet skulle foregå.

Deltagere ble rekruttert ved tilfeldig utvalg, men med følgende oppfylte kriterier: de måtte forstå norsk godt, studere medisin ved UiO på modul 6 eller 7, og de måtte ha tilgang til internett. Seleksjonen tok hensyn til studieretning og modul, men ikke til alder og kjønn. Vi rekrutterte medisinstudenter, både via mail, Facebook og direkte forespørsler, fra vårt eget kull (H13) og fra kullet under (V14). Vi rekrutterte 20 medisinstudenter (9 jenter og 11 gutter) og fordelte dem deretter tilfeldig ved trekning i to grupper (henholdsvis intervensjon GN og NG). Den ene gruppen (GN) ble tildelt den "gamle" e-læringsressursen fra 2004, mens den andre gruppen (NG) fikk den nye versjonen à 2016. De skulle ikke se på hele ressursen, men kun det ene leddet (kneet).

Deltakerne fikk videre instruks om at de skulle gå igjennom hele den tildelte internettleksjonen og deretter besvare quizen. De hadde altså på forhånd blitt tildelt enten den nye eller den gamle web-leksjonen som omhandlet *kneet*, avhengig av hvilken gruppe man ble randomisert til. Etter en viss tid (minimum 2 uker) skulle gruppe GN gjenta prosessen for den nye e-læringsressursen, mens gruppe NG skulle gjøre det samme for den gamle. Man skulle altså gjøre det samme med den ressursen man ikke hadde fått tildelt ved første gjennomgang. Tanken med å ha det adskilt i tid var at deltakerne forhåpentligvis hadde "glemt" noe av det de lærte i først runde. Dersom en student eksempelvis ble tildelt den gamle ressursen til å begynne med, skulle han/hun besvare quizen etter å ha sett på den gamle ressursen. Deretter skulle vedkommende gjennomgå den nye, for deretter igjen svare på

quizen. Til slutt skulle man besvare selve spørreundersøkelsen – som har som hovedmål å sammenlikne begge ressursene og belyse områder for forbedring.

Vi var ikke selv tilstede for å observere at quizen ble utført og at spørreundersøkelsen ble besvart, men dette ble avtalt på forhånd med de frivillige deltakerne. De fikk en gitt tidsramme de måtte fullføre undersøkelsene på, men de kunne selv bestemme hvor og når de ville gjøre det innenfor disse rammene. Spørsmålene ble sendt til deltakerne via mail, hvor de deretter skulle åpnes i Microsoft Word (eller tilsvarende), fylles inn, og sendes tilbake.

3.3.2 Anonymisering

Spørreskjemaene var i utgangspunktet ment å være anonyme. Men for at undersøkelsen skulle kunne gjennomføres på det tidspunkt og sted det passet deltakerne best, så var skjemaene av praktiske årsaker nødt til å sendes via mail eller Facebook. Vi forsikret deltakerne om at vi allikevel ville holde svarene anonyme.

3.3.3 Innholdet i undersøkelsene

Denne pilot-brukertesten ble i hovedsak utført for å kunne indentifisere brukerfeil i designet, for å sjekke om vi var på rett vei i revisjonsarbeidet og for å imøtekomme andre ønsker og nye ideer for dermed å kunne gjøre de andre web-leksjonene (som ikke er ferdige), enda bedre. Vi testet dette igjennom en faglig basert quiz i tillegg til en spørreundersøkelse. Mens quizen hadde som hensikt å fungere som et objektivt mål på læringsutbyttet, skulle spørreundersøkelsen dekke en mer subjektiv og kvalitativ del som omhandlet bruken og nytten av ressursene, og med mulighet for tilbakemeldinger og tanker om forbedringspotensial.

Quiz

Oppgavene i denne quizen, eller "mini-eksamenen", var ment for å teste hvor mye studentene hadde fått med seg av det faglige. Tanken var at den skulle fungere som et objektivt mål på hvor læringseffektiv den tildelte e-læringsressursen var. Spørsmålene skulle være kunnskapsmessig relevante, og de hadde som formål å bringe testpersonene innom de forbedringspunktene vi selv hadde formulert. Tanken bak dette var at vi skulle få sammenlignbare observasjoner på de to systemene (Vedlegg 1).

Spørreundersøkelse

Vi har ivaretatt oppsettet for spørreundersøkelsen som beskrevet ovenfor, ved å stille enkle bakgrunnsspørsmål til å begynne med, før vi involverer spørsmål om e-læringsressursene og deres kvalitet. Vi fant det nødvendig å dele spørreskjemaet inn i fire deler:

Del 1 Bakgrunn

Først blir pre-test spørsmål administrert. Disse inkluderer spørsmål om kjønn, alder, internettilgjengelighet og generell bruk av nettressurser.

Del 2 Sammenlikning

Versjonene av e-læringsressursen for henholdsvis 2004 og 2016 sammenliknes. Denne delen består av ikke-standardiserte spørsmål. De er formulert på en kvantitativ måte, med svaralternativer på en skala fra 1-5. Kjent ros og kritikk av gammel ressurs er forsøkt implementert.

Del 3 Evaluering av oppdatert e-læringsressurs

Her valgte vi å bruke et standardisert spørsmålsoppsett, System Usability Scale (SUS). SUS brukes som en metode for å måle den subjektive vurderingen av brukervennligheten i den nye ressursen ved hjelp av ti standardiserte spørsmål.

Del 4 Kvalitativ del

Avslutningsvis har vi en rekke kvalitative spørsmål som går mer ut på tilbakemeldinger, forbedringsforslag og fremtidig bruk. Deltakerne skal besvare disse med egne ord.

Det kreves en viss erfaring og forskningsinnsikt for å kunne stille gode spørsmål. Viktigheten av presise og riktig formulerte spørsmål er beskrevet over. For å unngå feil som kunne påvirke kvaliteten av undersøkelsen og i verste fall føre til feilaktige konklusjoner, fikk vi derfor god veiledning av både Per Grøttum og Lars Engebretsen i arbeidet vårt med å utforme spørsmålene (Vedlegg 1: Spørreskjema).

3.3.4. Evaluering og metodekritikk

For å gjennomføre evalueringen av internettressursen har vi valgt å fokusere på kun én av web-leksjonene, nemlig den som omhandler kneet. Dette skyldes både en tidsbegrensning i forhold til hva vi kan forvente at studentene setter av tid til for å besvare undersøkelsen vår, og det faktum at de andre web-leksjonene fortsatt er under revisjon (ferdigstilles ca 1. april 2018). Fordelen med å gjøre det slik er at vi lettere kan rekruttere forsøkspersoner, og ved å begrense mengden kan vi tenke oss til at det er mer sannsynlig at studentene setter seg ordentlig inn i det. En annen fordel er at vi samtidig har mulighet til å gjennomføre denne prosjektoppgaven parallelt som at vi arbeider med å oppdatere selve ressursen. Vi kan dermed ta med oss lærdom fra det vi får ut av spørreundersøkelsen om kneet, til å gjøre de andre leddene enda bedre. Ulempen er at studentene kun får sett på ett ledd, og dermed får et mindre sammenlikningsgrunnlag. En annen ulempe er at vi ikke får evaluert alle leddene og at vi da kan miste informasjon om at oppdateringen eventuelt ikke er like godt utført i alle web-leksjonene. Dette ivaretas imidlertid av at "kneet" også blir brukt som prototype i vårt arbeid med de andre leddene, slik at i alle fall oppbygningen og informasjonsmengden skal være omtrent den samme.

Av tidsmessige årsaker har vi valgt å utføre selve undersøkelsen over et halvt år, på medisinstudenter fra vårt eget kull (H13) og kullet under vårt (V14). Dette er en fordel med tanke på rekruttering både i forhold til at det er folk vi kjenner, og at studentene nå er på et tidspunkt i studiet hvor det er en god stund siden man hadde dette faget og at de selv ser det som gunstig med en repetisjon. En ulempe kan være at studentene har hatt faget før og dermed har en viss bakgrunnskunnskap, og at de som kjenner oss godt velger å delta, og at vi dermed får en viss type mennesker med i studien og følgende en "bias" i utvalget. Vi kunne

også ha inkludert andre målgrupper som undervisere, leger i turnus og ferdigutdannede leger i evalueringen, og evaluert over flere kull over en lengre tidsperiode. Tidsperioden mellom de to gjennomgangene for hver deltaker, ble satt til minimum to uker. Hadde vi hatt enda lengre tid i mellom så kunne vi sikret en enda lengre utvaskingsperiode, men vi var dessverre begrenset av tid og andre forhold som deltakernes eksamener o.l.

4 Resultater

4.1 Innhenting og Svarprosent

Web-leksjonen og spørreundersøkelsen til kneet ble fullstendig ferdigstilt i mai 2017. I den påfølgende måneden startet rekrutteringen av studentene på kull H13, som da gikk på modul 5 (MED4500). Kull H13 har en egen Facebook-gruppe som brukes jevnlig til å dele informasjon, og den består av 142 medlemmer. Her ble det lagt ut en fellesmelding som kort beskrev studien og hva det ville innebære å delta. Studentene ble bedt om å enten melde sin interesse under innlegget, eller kontakte oss direkte på mail eller via melding på mobil eller Facebook. Ni personer meldte sin interesse. Tre av disse sendte inn resultater allerede i juni, etter å ha gjennomgått første ressurs. Samtidig ble det sendt individuelle meldinger på Facebook til hvert medlem på kullet for å samle inn flere deltakere. Denne rekrutteringsprosessen tok ganske mye lengre tid enn ventet, til tross for iherdige forsøk med private meldinger til alle og en hver. Til slutt sa totalt 20 studenter seg villige til å delta. Alle disse studentene fikk deretter i september tilsendt link til den første ressursen de ble tilfeldig utvalgt til å starte med. Av ulike grunner var det kun 10 stykker (50 %) av disse som fremdeles hadde anledning til å delta. Vi innså at vi ble nødt til å inkludere kullet under også, for å få nok deltagere. Deretter startet en ny rekrutteringsprosess, hvor en annonse ble lagt ut på kullet under (V14), som da nettopp hadde begynt på modul 5 (MED4500). På kull V14 sin Facebook-side er det 267 medlemmer, og kun 6 av disse meldte sin interesse. Deretter ble det sendt individuelle meldinger til medlemmer på kull V14 til det var nok deltakere. Disse hadde dessverre ikke anledning til å starte intervensjonene før utført eksamen i november. I november ble altså alle de resterende deltakerne på tvers av både kull H13 og V14 informert om starten av intervensjonene, og det ble forsøkt å finne tidspunkt som passet for hver enkelt. Av ulike årsaker ble det ganske stor variasjon i starttidspunktet for første runde av intervensjonen for hver enkelt deltaker, men det ble allikevel forsøkt å finne et tidspunkt minst 2 uker senere til gjennomgang av andre runde av intervensjonen.

Det var totalt 409 medlemmer på kullsidene til H13 og V14 til sammen, som ble informert om studien via fellesmelding på kullsidene på Facebook. I den første rekrutteringen ble det sendt

ut meldinger til alle på kullsiden til H13, altså 142 stykker. Det ble totalt registrert 27 personer, altså 6.6 % av totalt antall studenter på kullene, som var interesserte i å delta i studien. Av disse gjensto det 20, altså 4.9 % som hadde anledning til å delta. Av disse 20 var det 11 menn og 9 kvinner. Gjennomsnittsalderen var på 27.4 år med et standardavvik på 6.2. Spredningen var fra 23 til 40 år.

I første runde ble 20 deltakere bedt om å gjennomgå den ene intervensjonen, utføre quiz og fylle inn spørreundersøkelse, og deretter sende disse resultatene inn. I andre runde, etter en periode på ca to uker skulle det samme gjentas men for den andre intervensjonen. Totalt 20 deltakere (100 %) hadde mulighet til å delta i første runde. I andre runde deltok 18 deltakere (90 %), og dette frafallet ble jevnt fordelt på intervensjonene ” gammel- før ny ressurs” (GN) og ”ny- før gammel ressurs” (NG). Altså var totalt antall resultater for intervensjon NG på 19/20 (95 %), og totalt antall resultater for intervensjon N var på 19/20 (95 %).

Videre følger resultatene til alle delene, som inkluderer *quizens* fem deler; Anamnese, Inspeksjon, Palpasjon, Bevegelighet og Spesielle tester; og *spørreundersøkelsen* som er delt opp i fire hoveddeler; 1 Bakgrunn, 2 Sammenlikning, 3 Evaluering av ny ressurs, og 4 Kvalitativ del.

4.2 Quiz

Quizen (se Referanse 8), bestod som nevnt av fem deler som henholdsvis inneholdt 7, 9, 5, 4 og 11 spørsmål. Hver deltaker ble bedt om å utføre quizen etter gjennomgang av sin tildelte e-læringsressurs. Etter en utvaskingsfase ble det gjort en ny gjennomgang av den andre e-læringsressursen, og quizen ble utført på nytt. Resultatene er samlet i Tabell 1.1 nedenfor.

Quiz

Deltaker

	1		2		3		4		5	
	G	N	G	N	G	N	G	N	G	N
Anamnese	43	57	71	86	71	71	71	100	43	43
Inspeksjon	67	78	89	89	67	78	89	67	78	89
Palpasjon	40	60	60	60	80	60	100	100	60	60
Bevegelighet	50	50	75	75	75	75	50	50	50	50
Spesielle tester	73	73	90	73	82	73	71	100	64	82
Sum	273	318	385	383	375	357	381	417	295	324

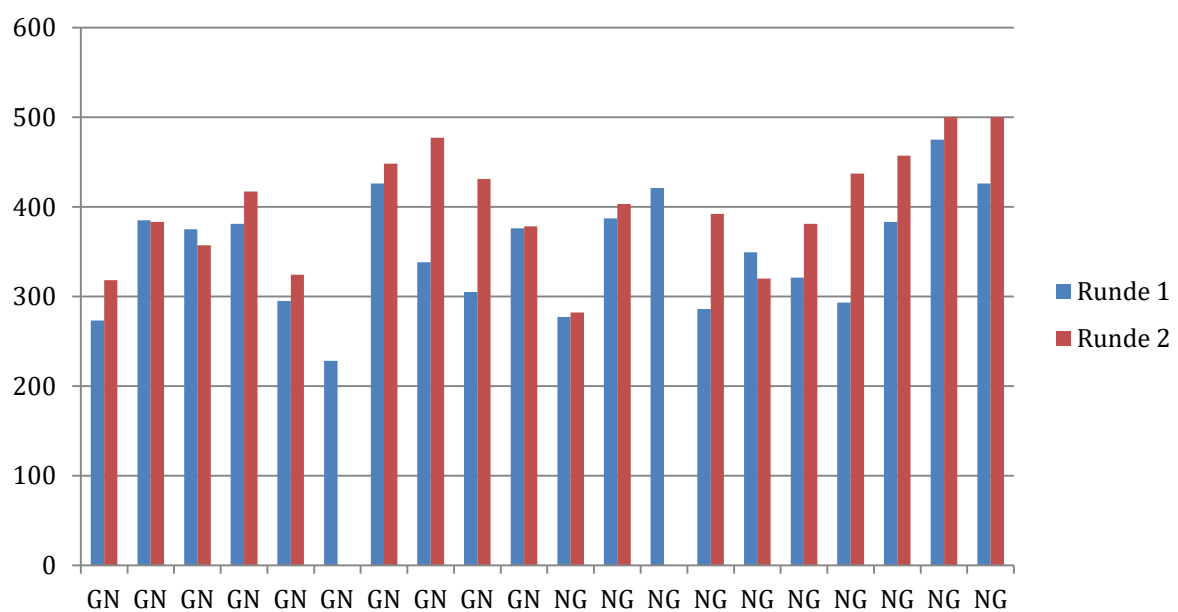
	6		7		8		9		10	
	G	N	G	N	G	N	G	N	G	N
Anamnese	35		86	86	57	86	71	17	86	86
Inspeksjon	45		78	100	78	100	56	89	67	78
Palpasjon	40		80	80	80	100	80	80	100	100
Bevegelighet	67		100	100	50	100	25	100	50	50
Spesielle tester	41		82	82	73	91	73	91	73	64
Sum	228		426	448	338	477	305	431	376	378

	11		12		13		14		15	
	N	G	N	G	N	G	N	G	N	G
Anamnese	43	43	57	57	85		57	57	43	57
Inspeksjon	67	56	100	89	89		89	78	89	50
Palpasjon	60	60	80	100	80		60	100	60	89
Bevegelighet	25	50	50	75	75		25	75	75	60
Spesielle tester	82	73	100	82	92		55	82	82	64
Sum	277	282	387	403	421		286	392	349	320

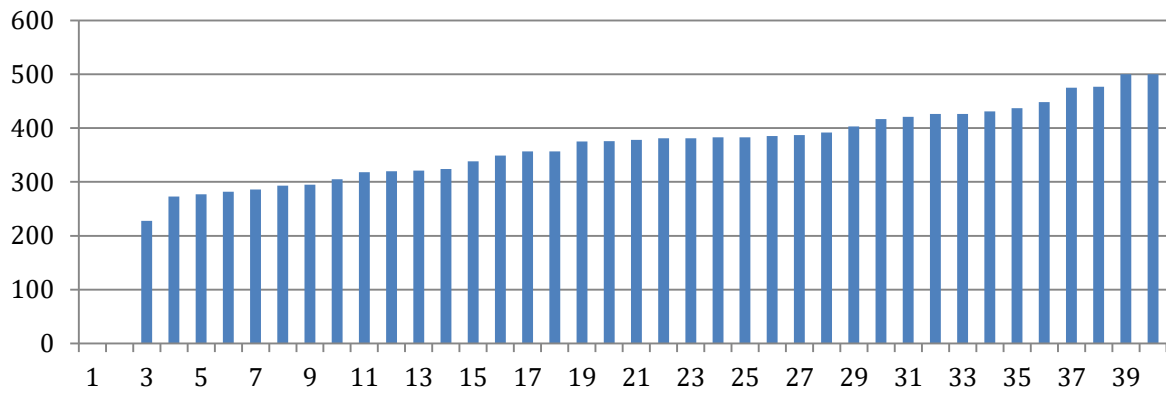
	16		17		18		19		20	
	N	G	N	G	N	G	N	G	N	G
Anamnese	56	57	57	86	57	86	100	100	71	100
Inspeksjon	67	78	89	89	89	100	100	100	100	100
Palpasjon	100	80	40	80	80	80	100	100	80	100
Bevegelighet	25	75	25	100	75	100	75	100	75	100
Spesielle tester	73	91	82	82	82	91	100	100	100	100
Sum	321	381	293	437	383	457	475	500	426	500

Tabell 1.1: Sammenlikning av quizresultater mellom gruppe GN (deltakere 1-10) og gruppe NG (deltakere 11-20). Venstre kolonne i hvert par viser resultater før overkryssing. Høyre kolonne viser resultater etter overkryssing.

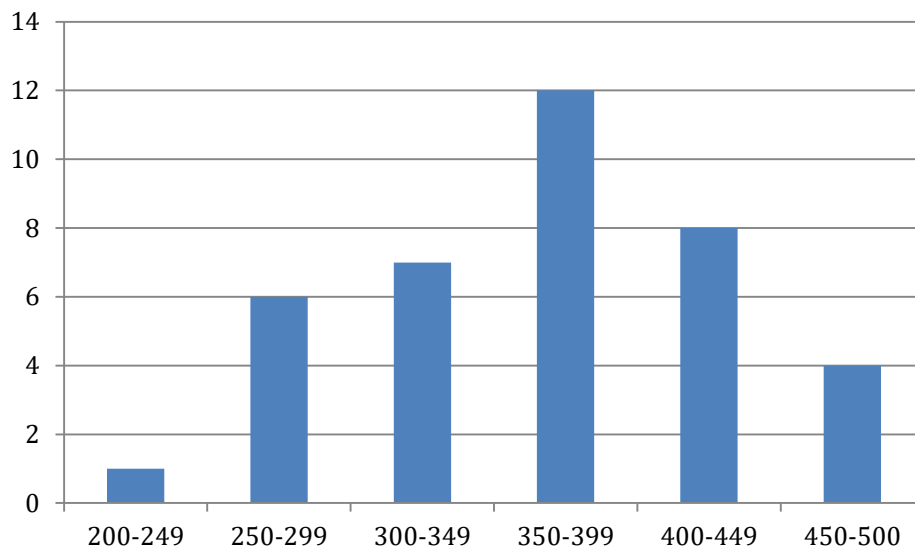
Tabell 1.1 viser resultatene fra quizen for både gruppe GN og gruppe NG, i første intervensjon til venstre og andre intervensjon til høyre. Deltakere 6 og 13 hadde kun anledning til å svare på første runde. Det er altså 19 resultater for test av den nye siden og 19 resultater for test av den gamle. Alle som først begynte på en quiz, utførte alle delene av quizen. Nederst under kategoriene er poengscorene summert, og denne har blitt brukt for å sammenligne resultatene.



Figur 2.1: Sumscore fra quiz for deltakere 1-20, skilles mellom Runde 1 (blå) og Runde 2 (rød) etter utvaskingsperiode.



Figur 2.2: Sumscore fra quiz for deltakere 1-20, sortert etter lavest score (frafall, 0) til høyest score (500).



Figur 2.3: Frekvensfordelingen av sumscore fra quiz (antall deltakere vs sumscore)

Fra resultatene kan vi se på spredningen i hver quizkategori, for å få et inntrykk om spredningen var større eller mindre i noen kategorier i forhold til andre. Spredningen fra lavest til høyest individuell *quizscore* på tvers av alle kategoriene er 75 (der maks poengsum er 100 og laveste oppnådde poengsum er 25). Spredning fra lavest til høyest individuelle *sumscore* (altså summen av de forskjellige delene av quizen for hvert individ) er på 372 (500-228). Spredningene fra lavest til høyest *quizscore* for alle deltakerne i de forskjellige kategoriene var som følger; Anamnese: 65 (100-35), Inspeksjon: 45 (100-55), Palpasjon: 60 (100-40), Bevegelighet: 75 (100-25), Spesielle tester: 59 (100-41).

Quizen inneholder faglige spørsmål som har som intensjon å objektivt kartlegge læringsutbyttet fra gammel og ny e-læringsressurs (heretter omtalt som intervensjon G og N). Resultatene fra Tabell 1.1 er oppsummert i Tabell 1.2. Gjennomsnittet av *sumscore* for alle quizer utført av alle studenter etter intervensjon G var 371.26 poeng ut av 500 mulige, med et gjennomsnitt på 74.5 prosentpoeng per kategori. Gjennomsnittet av *sumscore* for alle quizer utført av alle studenter etter intervensjon N var 382.12 poeng ut av 500 mulige, med et gjennomsnitt på 76.4 prosentpoeng per kategori.

Det er målt en forskjell i *sumscore* på 10.85, altså en gjennomsnittlig forbedring av quizresultater på 2.2 prosentpoeng per kategori. Det kan se ut som om den nye ressursen ga et bedre resultat enn den gamle. Vi kan da stille en hypotese om at intervensjon N er bedre enn intervensjon G. For å si med sikkerhet at dette resultatet ikke bare er grunnet tilfeldigheter, setter vi opp en nullhypotese H_0 om at det ikke er en forskjell mellom intervensjonene. Dersom denne nullhypotesen kan forkastes, kan vi si at det er en forskjell mellom intervensjonene.

Videre følger oppsettet for hypotesetestingen.

$$H_0: \mu_N = \mu_G$$

$$H_1: \mu_N \neq \mu_G$$

$$t = \frac{\mu_N - \mu_G}{\frac{\sigma_N^2}{n_N} + \frac{\sigma_G^2}{n_G}}$$

μ_G	μ_N	σ_G	σ_N	n_G	n_N	df	t
371.26	382.12	75.83	60.81	19	19	36	0.0218

Tabell 1.2: Analyse av quizresultater i Tabell 1.1. μ_G : gjennomsnittet av sumscore for alle quizer utført etter intervensjon G. μ_N : gjennomsnittet av sumscore for alle quizer utført etter intervensjon N.

Ved å slå opp i student t-distribusjonstabellen, finner man følgende nødvendige betingelse for å forkaste H_0 med et konfidensintervall på 95%:

$$df = 36 \rightarrow |t| > 1.688$$

$$|t| = 0.0218 < 1.688$$

På grunnlag av resultatene våres (som vist i tabell 1.2.) kan vi derfor ikke forkaste H_0 , og kan dermed ikke si med 95% sikkerhet at den nye e-læringsressursen ga bedre quizresultater enn den gamle.

4.3 Spørreundersøkelse

Spørreundersøkelsen (se Vedlegg 1) ble fylt ut av alle deltakerne etter de hadde gjennomgått begge e-læringsressursene. Spørreundersøkelsen ble som tidligere nevnt delt i 4 deler.

Del 1, *Bakgrunn*, stilte 4 spørsmål; 1.1.- 1.4. Spørsmål 1.3 og 1.4 skulle svares på en skala 1-5. Kort oppsummert var spørsmålene som følger: 1.1 *Kjønn*, 1.2 *Alder*, 1.3 *Tilgang til internett hjemme og på skolen*, 1.4 *Tidligere bruk av fakultetets e-læringsprogrammer*.

Del 2, *Sammenlikning*, stilte 9 hovedspørsmål; 2.1.-2.9, hvorav det ene, spørsmål 2.8, besto av 4 deler (2.8a - 2.8d), altså totalt 12 spørsmål. Alle disse spørsmålene skulle besvares på en skala 1-5, og alle spørsmålene skulle besvares om både ny og gammel ressurs. Kort oppsummert var spørsmålene som følger: 2.1 Samsvar med læreplanmål, 2.2 Relevans til turnus, 2.3 Samsvar med standardisert journaloppsett, 2.4 Oppdatering, 2.5 Systemhastighet, 2.6 For mye detaljer, 2.7 Oversiktlig, 2.8 Vurdering av innhold delt inn i 2.8a Illustrasjoner, 2.8b Skriftlig innhold, 2.8c Videoer, 2.8d Interaktivitet, og 2.9 Engasjerende.

Del 3, *Brukervennlighet i oppdatert ressurs*, stilte 10 spørsmål; 3.01-3.10. Alle disse skulle besvares på en skala 1-5, men disse spørsmålene skulle kun besvares om ny ressurs. Kort oppsummert var spørsmålene som følger: 3.01 Mersmak, 3.02 Unødvendig komplisert, 3.03 Lett å bruke, 3.04 Behov for hjelp til å bruke, 3.05 Sammenheng mellom deler, 3.06 Ulogisk system, 3.07 Lærer fort å bruke, 3.08 Vanskelig å bruke, 3.09 Føltes sikker, 3.10 Lærekrav før bruk.

Del 4, *Kvalitativ del*, stilte 7 spørsmål; 4.1.-4.7., hvorav alle utenom 4.3 skulle besvares med egne ord. Spørsmål 4.3 Tidsbruk skulle i utgangspunktet kun besvares om den nye ressursen, men deltakerne ble i etterkant også spurt om den gamle ressursen. De spørsmålene i spørreundersøkelsen som ga kvantitative svar, altså spørsmål 4.3 ble så samlet i Tabell 2.1a og Tabell 2.1b nedenfor.

Spørsmål

Deltaker

		1		2		3		4		5		6		7		8		9		10	
		G	N	G	N	G	N	G	N	G	N	G	N	G	N	G	N	G	N	G	N
1	1	K		M		M		K		K		K		M		M		K		K	
	2	26		24		23		24		24		26		23		26		23		40	
	3	5		5		5		5		4		4		5		5		5		5	
	4	3		3		3		3		3		2		2		4		3		3	
2	1	4	4	3	4	4	4	3	4			3		3	4	4	4	5	5	3	3
	2	4	4	2	4	3	4	1	2	2	4	2		3	4	3	5	4	5	4	3
	3	4	4	3	4	4	5	3	5	3	4	3		4	4	4	5	2	5	4	4
	4	4	4	3	4	3	5	1	3	2	5	2		3	4	4	4	1	5	3	4
	5	4	4	3	4	3	5	1	5	2	4	4		1	5	1	5	5	4	3	4
	6	4	3	5	3	4	4	4	2	4	2	4		2	1	4	2	2	2	2	2
	7	3	4	3	5	3	4	2	4	3	4	2		2	5	1	5	1	4	4	4
	8a	3	4	3	4	3	4	2	4	1	3	3		3	4	2	5	1	4	3	4
	8b	3	4	4	4	3	4	3	5	3	4	3		3	4	2	4	3	4	4	3
	8c	5	5	3	4	3	4	4	5	2	5	4		3	4	4	5	5	5	4	4
8d	3	5	3	4	3	4	1	3	2	4	3		2	3	1	3	2	2	2	3	
9	3	4	3	5	3	4	1	5	2	4	3		2	4	3	5	1	3	3	4	
3	1		4		4		4		4		2			3		3		4		5	
	2		1		3		2		2		2			2		1		1		1	
	3		5		4		4		4		5			5		5		5		5	
	4		1		1		1		1		1			1		1		1		1	
	5		5		3		4		4		5			4		5		5		5	
	6		1		2		2		2		1			3		1		1		4	
	7		5		4		4		5		5			5		5		5		5	
	8		1		2		2		1		1			1		1		1		1	
	9		5		4		2		4		5			5		5		5		5	
	10		1		2		1		1		1			1		1		1		1	
4	3	3.5	2	2.5	3	1.5	1.5		1	1.5	2	2		1.5	1	3	1	3	1.5	1.5	1

Tabell 2.1a: Resultater fra spørreundersøkelsen, gruppe GN (deltakere 1-10). Hvert tall i venstre kolonne korresponderer med tilhørende spørsmål, jfr. Vedlegg 1.

I Tabell 2.1a står resultatene fra deltakere 1-10, altså gruppe GN. Grundigere statistisk analyse av resultatene følger senere (se Tabell 2.2). Vi ønsker å belyse bakgrunnen til deltakerne i gruppe GN og NG. I første omgang beskrives derfor kun resultatene fra Del 1 separat for hver gruppe. Resultatene fra øvrige deler beskrives samlet under Tabell 2.2.

Del 1: Alle studiedeltakerne i gruppe GN besvarte alle spørsmålene. Her er en kort oppsummeringsresultatene er som følger.

- 1.1 *Kjønn*: av 10 deltakere i gruppe GN var det totalt 4 menn (40 %) og 6 kvinner (60 %).
- 1.2 *Alder*: gjennomsnittsalder på deltakerne var 25.90 år med et standardavvik på 5.11.
- 1.3 *Tilgang til internett hjemme og på skolen* : gjennomsnittet var 4.80 med et standardavvik på 0.42.
- 1.4 *Tidligere bruk av fakultetets e-læringsprogrammer*: gjennomsnittet var 2.90 med et standardavvik på 0.57.

Spørsmål 2.1 *Samsvar med læreplanmål* ble besvart i 17/20 (85 %) spørsmål, hvorav 9 om gammel ressurs og 8 om ny ressurs. Resterende spørsmål 2.2-2.9 ble besvart i 19/20 (95 %) spørsmål.

Del 3: Alle som leverte spørreundersøkelsen svarte på alle spørsmål i Del 3, men det var kun 9 deltakere (90 %) som leverte spørreundersøkelsen.

Del 4: kvantifiserte vi kun 4.3 Tidsbruk. Her fikk vi inn 18 resultater (90 %), hvorav 9 fra gammel og 9 fra ny ressurs.

Videre følger en kort oppsummering av resultatene fra gruppe NG (deltakere 11-20), deretter vil alle resultatene fra Del 1-4 analyseres nærmere, samlet.

Spørsmål

Resultater

		11		12		13		14		15		16		17		18		19		20			
		N	G	N	G	N	G	N	G	N	G	N	G	N	G	N	G	N	G	N	G		
1	1	M		K		M		M		M		M		K		M		M		K			
	2	23		23				33		27		26		27		23		23		40			
	3	5		5				5		5		5		5		5		5		5			
	4	3		3				3		1		4		3		2		3		3			
2	1	5	3	4	3							3	4	4	4	4	4	5	4	4	5	3	
	2	4	2	4	4			3	3	4	1	4	5	4	3	4	3	4	4	4	5	2	
	3	5	5	5	5			5	4	5	2	4	5	5	2	5	4	4	4	2	5	2	
	4	5	4	4	2			5	4	5	2	5	3	4	2	5	1	5	4	5	5	3	
	5	4	4	4	4			4	1	5	1	5	3	5	3	5	4	5	4	5	4	4	
	6	1	3	4	5			4	5	2	5	1	1	3	4	1	5	3	5	3	5	3	5
	7	5	2	4	2			3	2	5	1	3	4	4	2	5	3	5	3	5	3	5	2
	8a	5	2	3	3			5	3	5	4	3	3	3	4	5	3	5	3	5	3	5	2
	8b	5	3	4	3			4	3	4	2	2	3	3	2	3	3	5	4	5	2	5	2
	8c	5	5	4	3			4	4	5	3	4	3	4	2	5	2	5	3	5	2	5	2
	8d	3	3	3	3			2	2	3	1	3	3	4	1	4	2	5	2	5	2	5	2
8e	4	2	4	3			4	3	3	1	3	3	4	2	4	1	5	3	4	2	5	2	
3	1	4		4				4		2		3		3		4		5		5		5	
	2	1		3				2		2		2		1		1		2		1		1	
	3	5		4				4		5		5		5		5		5		5		5	
	4	1		1				1		1		1		1		1		1		1		1	
	5	5		3				4		5		4		5		5		5		5		5	
	6	1		2				2		1		3		1		1		1		1		4	
	7	5		4				5		5		5		5		5		5		5		5	
	8	1		2				1		1		1		1		1		1		1		1	
	9	5		4				4		5		5		5		5		5		5		5	
	10	1		2				1		1		1		1		1		1		1		1	
4	3	0.5	1	2	0.66			2.5	3			1.5	3	3.5	1.5	1.5	3	1	2	2	1.5	1	

Tabell 2.1b: Resultater fra spørreundersøkelsen, gruppe NG (deltakere 11-20). Hvert tall i venstre kolonne korresponderer med tilhørende spørsmål, jfr. Vedlegg 1.

I Tabell 2.1b står resultatene fra deltakere 11-20, altså gruppe NG. Grundigere statistisk analyse av resultatene følger senere. Nedenfor er resultatene kun fra gruppe GN oppsummert.

9/10 (90 %) deltakere i gruppe NG besvarte alle spørsmålene i Del 1. Deltaker 13 glemte å sende inn spørreundersøkelsen, og hadde ikke lenger mulighet til å delta. Deltakerens kjønn er allikevel notert. Totalt antall svar på spørsmål 1.1 er altså 10/10 (100 %), på spørsmål 1.2 var det 9/10 (90 %), på spørsmål 1.3 var det 9/10 (90 %) og på spørsmål 1.4 var det 9/10 (90 %). Resultatene er som følger.

- 1.1 *Kjønn*: av 10 deltakere i gruppe NG deltok totalt 7 menn (70 %) og 3 kvinner (30 %).
- 1.2 *Alder*: gjennomsnittsalder på deltakerne var 27.22 år med et standardavvik på 5.80.
- 1.3 *Tilgang til internett hjemme og på skolen* : gjennomsnittet var 5 med et standardavvik på 0 (alle svarte 5).
- 1.4 *Tidligere bruk av fakultetets e-læringsprogrammer*: gjennomsnittet var 2.78 med et standardavvik på 0.83.

Spørsmål 2.1 *Samsvar med læreplanmål* ble besvart i 15/20 (75 %) spørsmål, hvorav 8 om gammel ressurs og 7 om ny ressurs. Resterende spørsmål 2.2-2.9 ble besvart i 18/20 (90 %) spørsmål.

Alle som leverte spørreundersøkelsen svarte på alle spørsmål i Del 3, men det var kun 9 deltakere (90 %) som leverte spørreundersøkelsen.

I Del 4 kvantifiserte vi kun 4.3 Tidsbruk. Her fikk vi inn 17 resultater (85 %), hvorav 9 fra gammel og 8 fra ny ressurs..

Alle resultatene fra Tabell 2.1a og 2.1b ble samlet og prosessert, som vist i Tabell 2.2 nedenfor.

Spørsmål

Resultater

		n_M	n_K						
1	1	11	9						
		μ	σ	n					
	2	26.56	5.49	19					
	3	4.89	0.32	19					
	4	2.78	0.65	19					
		μ_G	μ_N	σ_G	σ_N	n_G	n_N	df	t
2	1	3.59	4.13	0.71	0.50	17	16	31	11.80
	2	2.89	3.94	1.10	0.73	19	18	35	10.88
	3	3.42	4.61	1.07	0.50	19	18	35	15.32
	4	2.68	4.50	1.06	0.62	19	18	35	21.80
	5	2.89	4.56	1.33	0.51	19	18	35	14.74
	6	3.84	2.39	1.26	1.04	19	18	35	-9.83
	7	2.37	4.33	0.90	0.69	19	18	35	27.81
	8a	2.68	4.17	0.82	0.79	19	18	35	20.68
	8b	2.95	3.94	0.62	0.80	19	18	35	17.43
	8c	3.37	4.56	1.01	0.51	19	18	35	16.63
8d	2.16	3.50	0.76	0.92	19	18	35	16.80	
	9	2.32	4.06	0.82	0.64	19	18	35	28.97
		μ_N	σ_N		n_N				
3	1	3.78	0.81		18				
	2	1.61	0.61		18				
	3	4.78	0.43		18				
	4	1.06	0.24		18				
	5	4.56	0.62		18				
	6	1.61	0.85		18				
	7	4.56	0.62		18				
	8	1.17	0.38		18				
	9	4.67	0.77		18				
	10	1.11	0.32		18				
		μ_G	μ_N	σ_G	σ_N	n_G	n_N	df	t
4	3	1.95	1.76	0.91	0.77	18.00	17.00	33.00	2.33

Tabell 2.2: Analyse av resultatene i Tabell 2.1a og Tabell 2.1b med relevante statistiske parametere.

Resultatene fra Del 1 i Tabell 2.2 inkludert svarprosent kan oppsummeres som følger:

- 1.1 Kjønn: 20/20 (100 %) resultater ble samlet inn for dette spørsmålet. Det var totalt 11 menn (55 %) og 9 kvinner (45 %) som deltok i spørreundersøkelsen.
- 1.2 Alder: 19/20 (100 %) resultater ble samlet inn for dette spørsmålet. Gjennomsnittsalderen var 26.56 år med et standardavvik på 5.49.
- 1.3 Tilgang til internett hjemme og på skolen: 19/20 (95 %) resultater ble samlet inn for dette spørsmålet. Gjennomsnittsscoren (1-5) var 4.89 med et standardavvik på 0.32.
- 1.4 Tidligere bruk av fakultetets e-læringsprogrammer: 19/20 (95 %) resultater ble samlet inn for dette spørsmålet. Gjennomsnittsscoren (1-5) var 2.78 med et standardavvik på 0.65.

Spørsmål 2.1-2.9 og 4.3 ble stilt ift både den nye og gamle ressursen, dermed ble gjennomsnittsverdiene for disse testet mot hverandre med tilhørende standardavvik.

		μ_G	μ_N	df	t	$t_{1-\frac{\alpha}{2},df}$	$ t - t_{1-\frac{\alpha}{2},df}$	<i>Konklusjon</i>
2	1	3.59	4.13	31	11.8	1.696	10.104	H_1
	2	2.89	3.94	35	10.88	1.69	9.19	H_1
	3	3.42	4.61	35	15.32	1.69	13.63	H_1
	4	2.68	4.50	35	21.8	1.69	20.11	H_1
	5	2.89	4.56	35	14.74	1.69	13.05	H_1
	6	3.84	2.39	35	-9.83	1.69	8.14	H_1
	7	2.37	4.33	35	27.81	1.69	26.12	H_1
	8a	2.68	4.17	35	20.68	1.69	18.99	H_1
	8b	2.95	3.94	35	17.43	1.69	15.74	H_1
	8c	3.37	4.56	35	16.63	1.69	14.94	H_1
8d	2.16	3.50	35	16.8	1.69	15.11	H_1	
9	2.32	4.06	35	28.97	1.69	27.28	H_1	
4	3	1.95	1.76	35	2.33	1.69	0.64	H_1

Tabell 2.3: Hypotesetesting av spørsmål 2.1-2.9 jfr Vedlegg 1.

Etter hypotesetesting med parett t-test og sammenlikning med student t-distribusjonstabellen finner vi t-verdier som vist i Tabell 2.3. Alle disse er større enn referanseverdien $t_{1-\frac{\alpha}{2},df}$, det vil si at $|t| - t_{1-\frac{\alpha}{2},df} > 0$ i alle tilfellene.

I alle tilfeller hvor i tillegg $t > 0$, kan vi konkludere (med mer enn 95% sikkerhet) med at de nye sidene hadde bedre resultater i spørreundersøkelsen enn de gamle sidene.

Ut ifra t-verdiene ser vi at alle tallene er positive med ett unntak, spørsmål 2.6.

Jfr. Vedlegg 1 ser vi at dette spørsmålet er ”Er det for mye detaljer?”, i dette tilfellet er 5 ”for mye detaljer” og 1 ”ikke for mye detaljer”. Her scoret altså de gamle sidene høyest.

- 2.1 Samsvar med læreplanmål: en t-verdi på 11.80 og df på 31, med et gjennomsnitt på 3.59 og 4.13 tyder på at den nye ressursen har signifikant bedre samsvar med læreplanmålene enn den gamle.
- 2.2 Relevans til turnus: en t-verdi på 11.29 og df på 35, med et gjennomsnitt på 2.89 (gammel) vs 3.94 (nye) tyder på at den nye ressursen er signifikant mer relevant til turnus enn den gamle.
- 2.3 Samsvar med standardisert journaloppsett: en t-verdi på 16.02 og df på 35, med et gjennomsnitt på 3.42 (gammel) vs 4.61 (nye) tyder på at den nye ressursen har signifikant bedre samsvar med standardisert journaloppsett enn den gamle.
- 2.4 Oppdatering: en t-verdi på 22.69 og df på 35, med et gjennomsnitt på 2.68 (gammel) vs 4.50 (nye) tyder på at den nye ressursen virker signifikant mer oppdatert enn den gamle.
- 2.5 Systemhastighet: en t-verdi på 15.45 og df på 35, med et gjennomsnitt på 2.89 (gammel) vs 4.56 (nye) tyder på at den nye ressursen er signifikant raskere å bruke enn den gamle.
- 2.6 For mye detaljer: en t-verdi på -10.45 og df på 35, med et gjennomsnitt på 3.84 (gammel) vs 2.39 (nye) tyder på at den nye ressursen, med statistisk signifikans, har mer rimelig detaljnivå, i forhold til den gamle.
- 2.7 Oversiktlig: en t-verdi på 28.76 og df på 35, med et gjennomsnitt på 2.37 (gammel) vs 4.33 (nye) tyder på at den nye ressursen er signifikant mer oversiktig enn den gamle.

- 2.8a Illustrasjoner: en t-verdi på 21.27 og df på 35, med et gjennomsnitt på 2.68 (gammel) vs 4.17 (nye) tyder på at den nye ressursen har signifikant bedre illustrasjoner enn den gamle.
- 2.8b Skriftlig innhold: en t-verdi på 17.78 og df på 35, med et gjennomsnitt på 2.95 (gammel) vs 3,94 (nye) tyder på at den nye ressursen har signifikant bedre skriftlig innhold enn den gamle.
- 2.8c Videoer: en t-verdi på 17.36 og df på 35, med et gjennomsnitt på 3.37 (gammel) vs 4.56 (nye) tyder på at den nye ressursen har signifikant bedre videoer enn den gamle.
- 2.8d Interaktivitet: en t-verdi på 17.17 og df på 35, med et gjennomsnitt på 2.16 (gammel) vs 3.50 (nye) tyder på at den nye ressursen er signifikant mer interaktiv enn den gamle.
- 2.9 Engasjement: en t-verdi på 29.95 og df på 35, med et gjennomsnitt på 2.32 (gammel) vs 4.06 (nye) tyder på at den nye ressursen er signifikant mer engasjerende enn den gamle.

Spørsmålene i Del 3 omhandler kun den nye ressursen. Blant de deltakere som svarte på Del 3 (18/20, 90 %), ble alle spørsmålene besvart.

- 3.01 Mersmak: Gjennomsnittsscoren (1-5) var 3.78 med et standardavvik på 0.81.
- 3.02 Unødvendig komplisert: Gjennomsnittsscoren (1-5) var 1.61 med et standardavvik på 0.61.
- 3.03 Lett å bruke: Gjennomsnittsscoren var 4.78 med et standardavvik på 0.43.
- 3.04 Behov for hjelp til å bruke: Gjennomsnittsscoren (1-5) var 10.6 med et standardavvik på 0.24.
- 3.05 Sammenheng mellom deler: Gjennomsnittsscoren (1-5) var 4.56 med et standardavvik på 0.62.
- 3.06 Ulogisk system: Gjennomsnittsscoren (1-5) var 1.61 med et standardavvik på 0.85.
- 3.07 Lærer fort å bruke: Gjennomsnittsscoren (1-5) var 4.56 med et standardavvik på 0.62.
- 3.08 Vanskelig å bruke: Gjennomsnittsscoren (1-5) var 1.17 med et standardavvik på 0.38.
- 3.09 Følte sikker: Gjennomsnittsscoren (1-5) var 4.67 med et standardavvik på 0.77.
- 3.10 Lærekraft før bruk: Gjennomsnittsscoren (1-5) var 1.11 med et standardavvik på 0.32.

Som nevnt tidligere hadde Del 4 *Kvalitativ del* kun ett kvantitativt spørsmål. Dette spørsmålet svarte 20/20 (100 %) deltakere på, med totalt 37/40 (92.5 %) resultater som ble innsamlet.

4.3 Tidsbruk: Gjennomsnittlig tidsbruk for den gamle ressursen var 1.95 timer med et standardavvik på 0.91. Gjennomsnittlig tidsbruk for den nye ressursen var 1.76 med et standardavvik på 0.77. Av tabell 2.3. som viser resultatene fra Spørsmål 4.3, ser vi at deltakerne brukte signifikant mindre tid på den nye ressursen enn den gamle.

Fordi resterende spørsmål i del 4 er av kvalitativ karakter er det ikke grunnlag for videre statistisk analyse av denne delen.

5 Diskusjon

Formålet med prosjektoppgaven er å evaluere e-læringsressursen ”klinisk undersøkelse av bevegelsesapparatet”. Resultatene fra spørreundersøkelsen viser en tydelig bedring på alle punkter av den nye ressursen i forhold til den gamle. Dette er vi svært fornøyde med. Det er allikevel noen punkter som kan gi opphav til feilkilder og som bør belyses.

5.1 Materiale

I denne undersøkelsen var det kun kne-delen av begge ressursene som ble evaluert. Det var flere årsaker til at vi valgte å gjøre det slik. For det første måtte vi ha en intervensjon som vi realistisk sett kunne rekruttere nok studenter til å gjennomføre. På grunn av tidsbruken ville det neppe la seg gjøre hvis deltakerne måtte gjennomgå flere enn ett ledd. Videre ønsket vi også å bruke kneet som prototype for de andre leddene, og derfor utnytte tilbakemeldingene fra denne studien før vi ferdiggjorde de andre leddenes e-lærings sider. Som et resultat av dette har vi ikke evaluert hele e-læringsressursen. Det kan selvfølgelig tenkes at akkurat kne-sidene er gjort dårligere i den gamle ressursen enn for eksempel rygg, og at det i den nye ressursen er motsatt – at kne er dårlig og et annet ledd er bra. Man kan derfor stille spørsmålet om hvorvidt man kan vurdere om materialet er representativt for resten av e-læringsressursen, siden vi kun evaluerer én av de åtte leksjonene. Dette forholdet kan være med på å svekke studiens eksterne validitet, som vil si at det kan være vanskelig å generalisere funnene fra et ledd til flere ledd. Hadde vi hatt anledning til å bruke alle leddene, hadde vi unngått denne utfordringen. Dette kan eventuelt undersøkes nærmere ved et senere tidspunkt når resten av e-læringsressursen er ferdigstilt. Vi har i midlertid forsøkt å ivareta en dette ved at kneet skal fungere som en prototype for de andre leddene. Tanken er at webleksjonene skal bli så like som mulig i både oppbygning og innhold, og at vi dermed forhindrer store forskjeller i kvaliteten på de ulike leksjonene.

Videre fungerte en rekke punkter om ønsker om forbedringer (som utdypet i utarbeidet av oss, som substans for både motivasjon til endringer i den nye versjonen, og som spørsmål i spørreundersøkelsen. Denne listen har vært vårt og noen av våre medstudenters inntrykk av

den gamle ressursen, og er følgelig en selvkomponert liste som kan være mangelfull. En annen måte å gjøre dette på, kunne ha vært å spørre flere studenter om hvilke endringer de ville foretrekke i en ny e-læringsressurs.

5.2 Studiedesign

Fordelen med en slik crossover-studie er at man ikke trenger å ha en egen kontrollgruppe, fordi deltakerne i de ulike gruppene fungerer som sine egne kontroller. Dette betyr også at vi ikke trengte å rekruttere et veldig stort antall studenter, som gjorde at vi faktisk kunne gjennomføre studien og samtidig få brukbare resultater. Valget om at hver deltaker skulle gjennomføre begge intervensjonene gjorde at hver og en også fikk et bedre sammenlikningsgrunnlag til å svare på spørsmål om de to ressursene. Dermed ble tilbakemeldingene deres om forbedringsforslag mer nyttig for oss. En annen fordel med vår studie var at den ga en stor inter-rater reliabilitet. Spesielt i resultatene til spørreundersøkelsen viste det seg å være godt samsvar mellom svarene til deltakerne. Dette artet seg gjennom små standardavvik og høye t-verdier, som taler for god statistisk signifikans.

I vår studie hadde deltakerne en "utvaskingsfase", en tidsperiode mellom de to intervensjonene. En stor utfordring ved denne måten å teste noe på, og dermed i crossover-studier generelt, er "carryover-effekten" mellom intervensjonene i hver gruppe. I vår studie ba vi altså deltakerne om å lære et stoff fra én ressurs, deretter vente to uker, og så lære stoffet på ny fra en annen ressurs. Carryover-effekten er her at deltakerne i forskjellig grad kan ta med seg kunnskapen fra den forrige intervensjonen videre, og at det dermed påvirker læringseffekten – som jo er noe av det vi ønsker å kartlegge (ved hjelp av quizen). Begrepsvaliditeten i studien vår er svekket av det faktum av at utførelse av quizen i seg selv ga en bedring av resultatet ved neste gjennomgang for nesten alle deltakerne. Dette kan bety at vi ender opp med å teste studentenes evne til å memorisere en quiz, istedet for å teste det faktiske læringsutbyttet til e-læringsressursene. Dette kunne vært unngått ved å bruke to forskjellige quizer. Det kunne vi ha laget, men da ville vi da hatt andre utfordringer med at disse kanskje ikke ville vært helt like i vanskelighetsgrad o.l. Dessuten ville det svekket den interne validiteten i studien, ettersom vi ville brukt to forskjellige måleverktøy, med svekket

reliabilitet. Det ville også vært interessant å se i fremtidige studier om en lenger utvaskingsperiode ville endret resultatene.

For å kunne si noe om overførbarhet i forhold til forbedringer av ressursen til en vanlig studiehverdag, er det flere punkter som kan kritiseres i forhold til prestasjon og tidsbruk i denne studien. For eksempel kan det være mindre insentiv til å prestere som deltaker i studien vår enn ved en vanlig eksamenssituasjon. Dessuten tar intervensjonene sted ved siden av studiet, altså er det "ekstra-arbeid" for deltakerne. Man kan spørre seg om deltakerne brukte like mye tid på intervensjonen som de ville brukt på en eksamensforberedelse. Det er også viktig å være bevisst på at vår studie kun har dokumentert effekten av hver intervensjon etter ca 2 timer gjennomsnittlig varighet. Vi kan dermed ikke generalisere disse resultatene i en mer realistisk eksamensperiodesetting, hvor man kunne forvente mange flere timer med innsats.

5.3 Resultater

5.3.1 Innhenting og Svarprosent

Optimalt sett ville det vært ønskelig å ha flere deltakere. Dette ville både gitt oss et mindre standardavvik og dermed høyere statistisk signifikans (som var en utfordring på Quiz-delen), og minsket effekten av eventuelle tilfeldige feil som kan oppstå av forskjellige årsaker. Av praktiske årsaker ville det vært lettere å gjennomføre studien med det ønskede antallet deltakere, hvis vi hadde noen ekstra deltakere i reserve som kunne steppet inn for de som av ulike grunner måtte trekke seg. Blant 20 deltakere vi endte opp med, var det kun 2 som ikke gjennomførte hele studien. Altså ble 38/40 intervensjoner gjennomført, noe som taler for god compliance og lavt frafall.

Den interne validiteten til studien vår var forholdsvis god fordi vi hadde en lav frafallsprosent kombinert med en høy svarprosent på stort sett alle spørsmålene i både quizen og spørreundersøkelsen. Den svekkes imidlertid noe av at det ble samlet inn 33/40 (82.5 %) resultater på spørsmål 2.1 (om samsvar med læreplanmålene), sammenliknet med 37/40 (92.5

%) på alle de andre oppgavene i Del 2 av spørreundersøkelsen. Dette var derimot det eneste spørsmålet som hadde en lavere svarprosent enn de andre spørsmålene.

Rekrutteringen var en stor utfordring i vår studie, fordi den krevde at studentene satt av en god del tid ved både første og andre intervusjon til å gjennomgå ressursene. Dette kan ha gitt en skjevfordeling (bias) ved at kun de særdeles motiverte deltok, for eksempel bekjente av undertegnede eller av veilederen. Både arbeidsmengden og tidsrammen påvirker dermed muligheten til å få et randomisert utvalg av befolkningen på kullene, som kan svekke studiens eksterne validitet. Fordi dette var en studie som tok tid og krevde mye av våre deltakere, ga det også en betydelig utfordring i forhold til innsamling av resultatene. Til senere gjennomføring av lignende studier på medisinstudenter kan det anbefales å sørge for at all innsamling av resultatene skjer tidlig i semesteret og i god tid før starten av eksamensperioden. Noen foretrakk også å få det i sommerferien, fordi de hadde for mye å gjøre med jobb og studier når semesteret var i gang.

Utvaskingsperioden var også vanskelig å standardisere på grunn av at deltakerne selv måtte få bestemte når intervusjonene skulle gjennomføres, i forhold til deres individuelle timeplaner.

Videre var det vanskelig å få deltakere til å overholde den perioden vi hadde fastsatt som utvaskingsperiode. Slik sett var denne utfordringen i compliance med på å svekke studiens interne validitet. Vi antar at compliance kunne blitt forbedret på flere nivå hvis studentene ville fått en eller annen form for kompensasjon, som for eksempel betaling, for deltakelsen

5.3.2 Quiz

Et av formålene med denne oppgaven var å se om innholdet i den nye ressursen var tilfredsstillende. Quizen er en del av den nye ressursen. Resultatene på denne viser at det verken var et påfallende antall deltakere som fikk lav sumscore, ei heller veldig høy sumscore, og svarprosentfordelingen tilsier dermed at quizen verken var for enkel eller for vanskelig. Vanskelighetsgraden på spørsmålene ville altså påvirke resultatene ved at vanskeligere spørsmål ville gitt et dårligere resultat, og omvendt, men resultatene viser altså at vi har truffet godt på det intenderte nivået. Dette kan tyde på at quizen kan fungere godt i den nye nettressursen.

Til tross for at resultatene viser at den nye ressursen gjennomsnittlig gjorde det bedre, var ikke denne forskjellen statistisk signifikant. Ut ifra formelen som ble brukt for å måle denne forskjellen, kan det være flere årsaker til dette. En av dem kan være at standardavviket er for stort. Standardavvik måler variasjonen mellom tallene i et datasett. Et for stort standardavvik tyder på at det er for stor variasjon i resultatene. Vi ser fra våre resultater at standardavvikene er nesten 5 ganger større enn forskjellen mellom sumscoren til de to ressursene. Det er flere måter å redusere denne variasjonen på. En måte ville ha vært å inkludere flere deltakere. Vi observerte i tillegg at antall spørsmål i hver kategori kan påvirke utfallet. For eksempel er det kun fire spørsmål i kategorien "Bevegelse". Dette kan fort resultere i veldig store forskjeller i score. Våre data viste en spredning på 75 (100-25). Det vi kunne ha gjort annerledes er å inkludere flere spørsmål i denne kategorien.

Alle faktorer som bidrar med å øke forskjellen på resultatene mellom individer eller mellom kategorier innen hvert individ, bidrar også til å øke standardavviket. Det er flere faktorer mellom individer, for eksempel tidsbruk. Personer som bruker mer tid vil mest sannsynlig gjøre det bedre på quizen, og omvendt. Dette gjelder både total tidsbruk for hele ressursen, så vel som tidsbruk innenfor de forskjellige kategoriene. Forhold kan være med på å påvirke tidsbruken enten bevisst eller ubevisst, kan være hvis vi i rekrutteringen for eksempel hadde sagt "sørg for å ta en ekstra grundig titt på spesielle tester, vi har mange flotte nye videoer der!", eller rykter om at for eksempel at anamnepunktet er veldig eksamensrelevant, kan muligens ha påvirket fordelingen av tidsbruken til deltakerne. Individuelle interesseområder påvirker selvsagt også hvilke deler man vier mer tid til. Disse faktorene som endrer

motivasjonen i de ulike delene av gjennomgangen, kan skape variasjon mellom forskjellige kategorier hos hvert enkelt individ.

Spørsmål 4.3 i spørreundersøkelsen var om hvor mye tid deltakerne brukte på den nye ressursen. Vi ser ut ifra dette at tidsbruk går fra 0.5 til 3 timer. Dette kan muligens påvirke hvor mye de lærer, og hvor bra de scorer. Dette er en tilfeldig feil i vår studie, og vil med nok deltakere ikke nødvendigvis endre gjennomsnittresultatet. Derimot kan det fremdeles ha en uheldig effekt på standardavviket og øke dette.

Videre kan forskjellige forventninger mellom individer påvirke motivasjonen deres for å bruke tid på ressursene. Hvis det er slik at de har en positiv forventning om at de nye ressursene er bedre, vil dette kunne føre til at de bruker mer tid på den nye ressursen. I tillegg vil kanskje en mer positiv forventning også påvirke i hvilken grad de husker stoffet som gjennomgås. Dette kan kategoriseres som en systematisk feil. En standardisert tidsramme hadde vært en løsning på denne utfordringen, og det ville vært interessant å se hvorvidt en slik standardisering ville påvirket resultatene. Det er selvfølgelig også mulig at deltakerne har forventninger i motsatt retning. En double-blind studiedesign ville trolig kunnet ha fjernet disse forventningene. Dessverre er ikke et slik design mulig, fordi studiedeltakerne allerede har gjennomgått de gamle sidene tidligere i studiet. I vår rekruttering har vi verken hatt mulighet til å pålegge en standardisert tidsramme som alle skal følge, ei heller kontrollere om en slik tidsramme faktisk ble fulgt.

Selve tidsfordelingen til hver enkelt deltaker, uavhengige av interesse og intensjon, kan også påvirke fordelingen av poengsummene og skape variasjon blant disse. Om en deltaker tok seg god tid til å gjennomgå ressursen grundig i starten, men fikk hastverk mot slutten, ville vedkommende eventuelt gjort det bedre på Anamnese og dårligere på Spesielle tester. Mange slike eksempler med ulik forventet effekt kan listes. De kategoriseres uansett som tilfeldige feil, som kan minimeres ved å ha flere deltakere.

Et annet punkt å belyse er om spørsmålene vi har laget er et godt mål på vurderingen av deltakernes læringsutbytte, altså hvor god den interne validiteten er. Det kom f.eks. fram i den kvalitative delen av spørreundersøkelsen at svaret på et av spørsmålene i quizen ikke kunne bli resonnert frem til med kunnskap fra den nye e-læringsressursen alene. Dette kan selvfølgelig ha en innvirkning på begrepsvaliditeten til studien vår.

En annen potensiell feilkilde for quiz-materialet er selve utformingen av spørsmålene. Hensikten med disse var å teste læreplanmålene i den nye revisjonen. Selv om den nye ressursen skal dekke læreplanmålene, kan quiz-spørsmålene ha blitt inspirert av det konkrete innholdet i den nye ressursen, til tross for at dette ikke var intensjonen. Dette kan igjen gi dem som har lest den nye ressursen en fordel når de skulle besvare disse fagrelaterte spørsmålene, og dermed påvirke resultatene i favør av den nye ressursen.

Som nevnt over var utfordringen med å lage en slik quiz å lage gode svaralternativer. Om de var for lette kunne svaret bli for opplagt om motsatt, og vi ville dermed ikke få en god sammenlikning av de to ressursene. Alternativet vårt var å lage quizen om til en frisvarsoppgave, men da ville problemet blitt hvordan vi objektivt skulle kunne bedømme svarene. Ved en frisvarsoppgave kan det være mange løsninger avhengig av hvordan man tolker spørsmålet, og mange svaralternativer kan være riktige. Vi valgte derfor å lage en quiz med et forsøk på å lage middels vanskelige spørsmål, og vi hadde med svaralternativer for å kunne kvantifisere og objektivisere resultatene så presist som mulig.

I og med at vi ikke hadde en online registrering av quizresultatene direkte, kan vi i prinsippet heller ikke være sikre på at studenter ikke har ”jukset” eller at de ved en feiltakelse har sendt inn et annet resultat enn de faktisk fikk. Dette kan være en kilde til tilfeldige feil.

5.3.3 Spørreundersøkelse

Bakgrunn

Vi vet i utgangspunktet at alle våre deltakere er medisinstudenter som har kommet omtrent like langt i studiet. Dette er med på å styrke studiens eksterne validitet. Vi ønsket derimot videre å kartlegge bakgrunnsinformasjon for å kunne avgjøre i større grad om utvalget var representativt. Kjønnfordelingen var 11 gutter og 9 jenter. Kjønnfordelingen på kullene derimot var ved studiestart henholdsvis 62 % kvinner på H13 og 75.5 % kvinner på V14, med ca samme antall studenter på kullene nå som den gang. Det er uvisst hva den eksakte fordelingen er i dag, men vi antar at kjønnfordelingen er noenlunde lik i dag. Det vi ser er at disse to kjønnfordelingene er ganske ulike kjønnfordelingen i vår studie, hvor vi fant en overrepresentasjon av menn. Det kan være grunner til dette. Det kan for eksempel tenkes at det mannlige kjønn er mer interessert i ortopedi og idrettsmedisin enn kvinner, og at de derfor har meldt seg frivillig til å delta i større grad. Gjennomsnittsalderen til deltakerne var 26.56 med et standardavvik på 5.49, Gjennomsnittsalderen på kullene har vi ikke gått videre innpå.

Videre viser resultatene at studentene hadde svært god internettilgang hjemme og/eller på Universitetet. Det finnes ingen mobilversjon av e-læringsressursen per dags dato, men internettsiden er fullt mulig å bruke på mobilen også. God nettilgang er viktig fordi det fører til at filmene lastes raskt ned, og tilgjengeligheten gjør at alle studentene har mulighet til å gjennomføre oppgavene. Dette er en forutsetning for at resultatene i spørreundersøkelsen vår skal være representative.

Resultatene viser at på en skala fra 1-5 var gjennomsnittsbruket av UiO sine e-læringsressurser 2.78 med et standardavvik på 0.65. Dette kan tolkes som at e-læringsressursene generelt ikke ser ut til å ha en stor plass i tilegningen av ny kunnskap, til tross for at internettilgangen er så god, men at den fortsatt spiller en viktig rolle. Det kan være flere grunner til at den ikke scorer høyere på bruken. En faktor kan være at man ikke vet at ressursene eksisterer, eller at man har erfaring med at de ikke er særlig gode. Et annet poeng kan være at det allerede finnes gode filmer på YouTube, eller at folk foretrekker bøker og forelesningsmateriale. Vårt inntrykk er imidlertid at de fleste ikke vet så mye om UiO sine

egne nettressurser. Bedre promotering fra forelesere er et mulig virkemiddel her, og særlig dersom foreleserne selv har sett dem og kan gi dem et kvalitetsstempel. Det er veldig variabel kvalitet på de ulike medisinske fagsidene. I vårt tilfelle var den gamle ressursen som omhandlet bevegelsesapparatet ganske tung og uoversiktlig, og egnet seg for eksempel ikke særlig godt til en kort eksamensrepetisjon. Vi mener at e-læringen har et stort potensiale i forhold til læring av ny kunnskap fordi den både inneholder skriftlig materiale, visuelle videoer, og i tillegg gir et innblikk i hvordan praksisen foregår lokalt her i Norge. Det blir en skreddersydd "oppskrift" spesielt rettet mot norske medisinstudenter og leger i turnus, og burde absolutt få mer oppmerksomhet. Dette forutsetter imidlertid at ressursene er godt lagd og er oppdaterte, som jo er vårt mål med denne fornyelsen av web-leksjonene.

Sammenlikning

For å kunne vurdere om den nye ressursen faktisk var bedre enn den gamle, valgte vi å ha med en egen del med sammenlikning av de to ressursene. Vi mente at den gamle ressursen hadde forbedringspotensialer på flere områder, som nevnt over i diskusjonen under "Materiale". Denne sammenlikningen inneholdt spørsmål som hovedsakelig hadde to formål. Det første var å avklare hvorvidt disse områdene faktisk trengte forbedring, eller om andre studenter var fornøyde med dem slik de var. Det andre var å finne ut om vi faktisk hadde klart å forbedre dem i vårt arbeid med den nye ressursen. Resultatene viser en statistisk signifikant forskjell (med et 95 % konfidensintervall) i favør av ny ressurs på alle punkter. Det vil si at vi tilsynelatende har klart å forbedre den gamle ressursen. De viser også i utgangspunktet lave gjennomsnittsscorer for den gamle ressursen.

Det er mange faktorer som kan ha påvirket resultatene fra spørreundersøkelsen i denne sammenlikningsdelen. Det er for eksempel nærliggende å tenke at deltakerne har en forventning om at den nye ressursen er bedre, og derfor har en tendens til å gi den en litt bedre score. Dette svekker studiens interne validitet, og kan gjøre det vanskeligere å trekke en gyldig konklusjon ut ifra resultatene våre. En annen utfordring ved en slik spørreundersøkelse er at kan være utfordrende å sørge for at svarene på de forskjellige spørsmålene er uavhengige av hverandre, og at de ikke alle har blitt påvirket av en felles feilkilde (konfundering). For eksempel rapporterte en del av studiedeltakerne at de slet seg igjennom stoffet mot sin vilje ved gjennomgang av de gamle sidene. Det er mulig at denne

generelle følelsen ga en priming-effekt og påvirket de ulike spørsmålene i denne delen i negativ retning, som kan være en kilde til systematisk feil.

Brukervennlighet i oppdatert ressurs

System Usability Scale er en fordel fordi det er en rask og anvendbar måte å kvantifisere informasjon som i utgangspunktet er kvalitativ. En utfordring er midlertid å sørge for at denne kvantifiseringen nøyaktig gjenspeiler det den beskriver (intern validitet), samtidig som det er enighet mellom deltakerne om hva som ligger i denne kvantifiseringen (inter-rater reliabilitet). Derimot er vi ikke ute etter objektive vurderinger av f.eks hva det vil si å være en "vanskelig" ressurs, men heller hvordan deltakerne selv opplever ressursen. Dette er SUS godt egnet til å vurdere. I en studie om evaluering av brukervennlighet i nettsider (Tullis & Stetson, 2004) (10) ga SUS resultater med høyest reliabilitet på tvers av utvalgsstørrelser, blant flere spørreundersøkelser undersøkt. For de spesifikke forholdene undersøkt i deres studie, var 12-14 deltakere nok til å gi resultater med god realibilitet (10). Disse funnene er i tråd med vår egen studie, og sånn sett kan vi si at denne skalaen et godt supplement til de andre måleverktøyene vi brukte.

Resultatene fra Del 3 viste tydelige forskjeller mellom de ulike spørsmålene, med relativt små standardavvik. Denne relativt gode inter-rater reliabiliteten styrker troverdigheten av funnene våre. Det var en svarprosent på 18/20 (90 %), pga 2 frafall fra studien som en helhet, det vil si at alle som svarte på spørreundersøkelsen om den nye ressursen, også svarte på Del 3.

Kvalitativ del

Resultatene fra Spørsmål 4.3 viser at deltakerne brukte signifikant mindre tid på den nye ressursen enn den gamle. Dette er et lovende resultat, gitt at det ikke kunne vises en signifikant forskjell i sumscore fra quizen. Forskjellen i tidsbruk kan tyde på at den nye ressursen er mer tidseffektiv og kommer frem til poenget fortere, noe som selvfølgelig er viktig både for studenter og leger.

Den kvalitative delen var i prinsippet ikke del av problemstillingen til denne oppgaven, men vi inkluderte den allikevel for å samle informasjon om hvordan vi kan bedre både denne nye ressursen for kne, men også de andre leddene som ikke er ferdigstilte ennå.

Det var en stor variasjon i svarene som de ulike deltakerne oppga, men som sagt var tilbakemeldingene i hovedsak veldig gode. En kort oppsummering av resultatene viser at studentene er svært fornøyde med følgende: ressursen var ryddigere og mer navigerbar. Det var langt flere og bedre illustrasjoner. Videoene var kortere og lettere å slå opp i. Det var mindre tidkrevende. Det var godt organisert, ”kort og konsist”, slik at en eventuell eksamensrepetisjon ville bli lettere. De likte også at alt innenfor et tema var på samme side, noe som gjorde det lettere å navigere og få oversikt.

Blant forbedringsforslag kom det varierende kommentarer hovedsakelig rettet mot det skriftlige innholdet og illustrasjonene. De kunne tenke seg flere kasus, mer presise og konsise setninger, og bedre forklaring av illustrasjonene ved hjelp av piler og tekst. Noen kommenterte også at vi gjerne kunne gjøre quizene mer omfattende, spesielt under enmet Bevegelighet. Vi har selv notert at denne quizen består av 4 spørsmål, som svekker dens evne til å gi et godt bilde av kunnskapsnivået til hver deltaker.

Alt i alt er vi svært fornøyde med tilbakemeldingene – både negative og positive – som vi fikk i den kvalitative delen.

6 Konklusjon

I en periode fra juni 2017 til januar 2018 gjennomførte vi en crossover-studie og en spørreundersøkelse blant en gruppe medisinstudenter der målet var å belyse problemstillingen i vår prosjektoppgave som var som følger: "Har vi klart å forbedre det som trengte forbedring, og å ivareta det som var bra, i arbeidet med oppdateringen av e-læringsressursen "Klinisk undersøkelse av bevegelsesapparatet"?".

Ut ifra resultatene fra sammenlikningsdelen av spørreundersøkelsen kan vi med statistisk signifikans, med et 95 % konfidensintervall, konkludere med at den nye siden har bedre samsvar med læreplanmål, er mer relevant til turnus, har bedre samsvar med standardisert journaloppsett, virker mer oppdatert, er raskere å bruke og har et mer rimelig detaljnivå. Den er mer oversiktlig, har bedre illustrasjoner og videoer, er mer interaktiv og mer engasjerende enn den gamle ressursen. I tillegg viser resultatene at deltakerne brukte signifikant mindre tid på den nye ressursen i forhold til den gamle, uten at det tilsynelatende påvirket hvordan de presterte faglig. Resultatene fra quizen viste i midlertid at gjennomsnittssummen var høyere etter gjennomgang av den nye ressursen enn den gamle, men denne forskjellen var ikke statistisk signifikant.

Videre viste resultatene fra Del 3 av spørreundersøkelsen at det var bred enighet blant deltakerne om at de godt kunne tenke seg å bruke ressursen ofte. De syntes ikke at den var unødvendig komplisert, men tvert imot lett og brukervennlig - som var et av våre hovedmål. Dette tyder på at vår nye e-læringsressurs kan ha gjort studentene mer positive til å benytte seg av fakultetes e-læringsressurser, og at dette er et område med mye potensiale som det burde satses mer på i fremtiden. I tillegg til å ha fått statistisk signifikante resultater som nevnt ovenfor, har vi i den kvalitative delen fått gjennomgående positive tilbakemeldinger på den nye e-læringsressursen. Vi kan med dette konkludere med at vi har klart å forbedre det som trengte forbedring, og ivareta det som fungerte godt, i arbeidet med oppdateringen av e-læringsressursen "Klinisk undersøkelse av bevegelsesapparatet".

7 Referanser

1: Bruaset, Haugen, Kristoffersen, Strand

Evaluering av webleksjonene “Klinisk undersøkelse av bevegelsesapparatet”.

Prosjektoppgave (2005).

2: D. Cook, A. Levinson, S. Garside

Internet-based learning in the health professions

Journal of American Medical Association, 300 (10) (2008), pp. 1181–1196

3: D. Cook, S. Garside, A. Levinson, D. Dupras, V. Montori

What do we mean by web-based learning? A systematic review of the variability of interventions

Medical Education, 44 (2010), pp. 765–774

4: D. Cook, A. Levinson, S. Garside

Instructional design variations in internet-based learning for health profession education: a systematic review and meta-analysis

Academic Medicine, 85 (5) (2010), pp. 909–922

5: <http://tidsskriftet.no/2012/04/sprakspalten/overkrysningsstudie-fremfor-kryssforsok>

6: Friis S, Vaglum P. Fra ide til prosjekt: En innføring i klinisk forskning. 2. utg. Oslo: Tano Aschehoug; 1999.

7: <http://research-methodology.net/research-methods/survey-method/questionnaires-2/>

8: Ny e-læringsressurs, intervensjon N

<http://meddev.uio.no/elaring/fag/ortopedi/kne/us.shtml#>

9: Gammel e-læringsressurs, intervensjon G

<http://meddev.uio.no/elaring/lcms/undersokelse/kne/>

10: T. S. Tullis, J. N. Stetson, A Comparison of Questionnaires for Assessing Website Usability 2004 Available from:

https://www.researchgate.net/publication/228609327_A_Comparison_of_Questionnaires_for_Assessing_Website_Usability [accessed Jan 22 2018].

III VEDLEGG

Vedlegg 1 Spørreskjema

SPØRSMÅL TIL SPØRREUNDERSØKELSEN:

1) BAKGRUNN:

- 1.1 Kjønn (Kvinne/mann)
- 1.2 Alder (skriv ned selv eller rullgardin)
- 1.3 Tilgjengelighet (internett-tilgang hjemme/skole) (1-5)
- 1.4 Bruk av IT: hvor mye bruker du fakultetets e-læringsprogrammer (1-5)

2) SAMMENLIKNING:

	GAMMEL:					NY:				
	1	2	3	4	5	1	2	3	4	5
2.1 I hvilken grad imøtekommer ressursen læreplanmålene på medisinstudiet?										
2.2 Hvor tydelig kommer det fram hva du som turnuslege skal gjøre i klinikken?										
2.3 Hvor logisk er oppsettet i forhold til det standardiserte journalkonseptet?										
2.4 Virker materialet oppdatert?										
2.5 Er systemet raskt? (Her menes det tekniske, som hastigheten på nedlastning av videoer ol.)										
2.6 Er det for mye detaljer?										
2.7 Hvor lett er det å få oversikt over det du skal lære?										
2.8 Hvor fornøyd er du med										
2.8a illustrasjonene?										

2.8b	det skriftlige innholdet?		
2.8c	videoene?		
2.8d	mulighetene for interaktivitet?		
2.9	Synes du ressursen er engasjerende (vekker den din interesse for bevegelsesapparatet?)		

3) Brukervennlighet i OPPDATERT RESSURS:

Vennligst sett kryss i kun en rute pr. spørsmål (1-5, der 1 er sterkt uenig, og 5 er sterkt enig)

- | | |
|------|---|
| 3.01 | Jeg kune tenke meg å bruke dette systemet ofte |
| 3.02 | Jeg synes systemet var unødvndig komplisert |
| 3.03 | Jeg synes systemet var lett å bruke |
| 3.04 | Jeg tror jeg vil måtte trenge hjelp fra en person med teknisk kunnskap for å kunne bruke dette systemet |
| 3.05 | Jeg synes at de forskjellige delene av systemet hang godt sammen |
| 3.06 | Jeg synes det var for mye inkonsistens i systemet. (Det virket "ulogisk") |
| 3.07 | Je vil anta at folk flest kan lære seg dette systemet veldig raskt |
| 3.08 | Jeg synes systemet var veldig vanskelig å bruke |
| 3.09 | Jeg følte meg sikker da jeg brukte systemet |
| 3.10 | Jeg trenger å lære meg mye før jeg kan komme i gang med å bruke dette systemet på egen hånd |

4) Kvalitativ del:

- 4.1 Hva var nyttig i den nye ressursen?
- 4.2 Hva var ikke nyttig?
- 4.3 Hva savnet du i den nye ressursen?
- 4.4 Hvor mye tid brukte du på å gå igjennom "Kneet" i ny ressurs
- 4.5 Forslag til forbedringer eller andre endringer?
- 4.6 Synes du vi burde ha mer eller mindre av noe?
- 4.7 Tror du at du vil bruke den nye e-BA? I så fall hvorfor/hvorfor ikke?

Vedlegg 2 Læreplanmålene

Fag: Kirurgi: Ortopedi

Modul 3 UE: 1 hele

Læringsmål

Etter fullført modul skal studentene kunne ta opp anamnese, utføre en klinisk undersøkelse og kjenne til de vanligste ortopediske sykdommer og tilstander i bevegelsesapparatet.

Læringsutbytte

Kunnskaper

Etter fullført modul skal studentene kunne:

- gjøre rede for symptomer og funn ved de vanligste sykdommer, brudd og skader i bevegelsesapparatet
- gjøre rede for avvik fra normalfunksjonen i nakke, rygg, skulder, albue, arm, hånd, bekken, hofta, kne, ankel og fot
- gjøre rede for bruk og indikasjon av supplerende undersøkelser ved tilstander i bevegelsesapparatet
- gjøre rede for forebygging, behandling og rehabilitering av de vanligste sykdommene i bevegelsesapparatet

Ferdigheter

Etter fullført modul skal studentene kunne:

- gjennomføre målrettet anamneseopptak med vekt på å kartlegge funksjon i bevegelsesapparatet
- gjennomføre en klinisk undersøkelse rettet mot sykdommer og brudd i rygg og nakke, skulder, albue og hånd, hofta, bekken, kne, ankel og fot
- gjennomføre diagnostiske vurderinger og foreslå videre utredninger av barn og voksne
- drøfte fordeler og ulemper ved kirurgisk behandling av tilstander og sykdom i bevegelsesapparatet

Fag: Kirurgi: Ortopedi

Modul 8 UE: 1 hele

Blokk: ortopedi/nevrokirurgi/ revmatologi

Læringsmål

Etter fullført modul 8 skal studentene ha tilegnet seg kunnskap om indikasjon for, og effekt av, ortopedisk kirurgi og rehabilitering. Studentene skal videre kunne samarbeide tverrfaglig innenfor et behandlingsteam og mellom instanser.

Læringsutbytte

Kunnskaper

Etter fullført modul 8 skal studentene kunne gjøre rede for diagnostikk og behandling av:

- skader, frakturer, infeksjoner og degenerative sykdommer av rygg og nakke, skulder, overarm, albue, håndledd og hånd, hofte, bekken, lår, kne, legg, ankel og fot
- benigne og maligne tumores i bein, brusk og muskel
- medfødte tilstander i barneortopedi som klumpfot, metatarsus varus, vertikal talus, medfødt hoftedysplasi, Legg-Calvé-Perthes sykdom og juvenil hofte-epifysiolyse
- multitraumer og traumatiske blødninger (bekkenfrakturer og frakturer i store bein)
- osteoporose med patologiske brudd

Ferdigheter

Etter fullført modul 8 skal studentene kunne:

- injeksjonsbehandle og tapping av ledd
- anvende lokalbedøvelse og ledningsblokade
- utføre sårbehandling
- frakturbehandle ved reponering, gipsing og annen immobilisering
- vurdere henvisning til spesialisthelsetjenesten og hvilke pasienter som trenger øyeblikkelig hjelp innleggelse i ortopedisk avdeling

Generell kompetanse

Etter fullført modul skal studentene kunne:

- samhandle med andre fagpersoner i håndteringen av pasienter med ortopediske lidelser